

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2002, 08:01:57 ; Search time 70.27 Seconds  
(without alignments)  
1824.686 Million cell updates/sec

**Title:** US-09-680-514-4 COPY 526 1047

perfect score: 522  
Sequence: 1 GGGCCAACATATCGGGCTC.....TACGCCACCTTGCCAGCCC 522

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45

```
Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	486.8	93.3	525	1	US-08-434-411-1	Sequence 1, Appli	
2	486.8	93.3	525	1	US-08-434-402-1	Sequence 1, Appli	
3	486.8	93.3	525	1	US-08-783-288-1	Sequence 1, Appli	
4	486.8	93.3	525	2	US-08-890-640-1	Sequence 1, Appli	
5	486.8	93.3	525	6	5194592-25	Patent No. 5194592	
6	474.6	90.9	546	3	US-08-469-318-177	Sequence 177, App	
7	474.6	90.9	546	3	US-08-469-318-178	Sequence 178, App	
8	474.6	90.9	546	3	US-08-468-609A-177	Sequence 177, App	
9	474.6	90.9	546	3	US-08-468-609A-178	Sequence 178, App	
10	474.6	90.9	546	5	PCT-US95-01185-177	Sequence 177, App	
11	474.6	90.9	546	5	PCT-US95-01185-178	Sequence 178, App	
12	474.6	90.9	921	3	US-08-469-318-72	Sequence 72, Appl	
13	474.6	90.9	921	3	US-08-469-318-75	Sequence 75, Appl	
14	474.6	90.9	921	3	US-08-469-318-78	Sequence 78, Appl	
15	474.6	90.9	921	3	US-08-469-318-84	Sequence 84, Appl	
16	474.6	90.9	921	3	US-08-468-609A-72	Sequence 72, Appl	
17	474.6	90.9	921	3	US-08-468-609A-75	Sequence 75, Appl	
18	474.6	90.9	921	3	US-08-468-609A-78	Sequence 78, Appl	
19	474.6	90.9	921	3	US-08-468-609A-84	Sequence 84, Appl	
20	474.6	90.9	921	5	PCT-US95-01185-72	Sequence 72, Appl	
21	474.6	90.9	921	5	PCT-US95-01185-75	Sequence 75, Appl	
22	474.6	90.9	921	5	PCT-US95-01185-78	Sequence 78, Appl	
23	474.6	90.9	921	5	PCT-US95-01185-84	Sequence 84, Appl	
24	474.6	90.9	966	3	US-08-469-318-68	Sequence 68, Appl	
25	474.6	90.9	966	3	US-08-469-318-70	Sequence 70, Appl	
26	474.6	90.9	966	3	US-08-469-318-71	Sequence 71, Appl	
27	474.6	90.9	966	3	US-08-469-318-73	Sequence 73, Appl	

## ALIGNMENTS

```

1
RESULT 1
US-08-434-411-1
; Sequence 1, Application US/08434411
; Patent No. 5681720
; GENERAL INFORMATION:
; APPLICANT: KUGA, TETSURO
; APPLICANT: MIYAJI, HIROMASA
; APPLICANT: SATO, MORIYUKI
; APPLICANT: OKABE, MASAMI
; APPLICANT: MORIMOTO, MAKOTO
; APPLICANT: ITOH, SEIGA
; APPLICANT: YAMASAKI, MOTOO
; APPLICANT: YOKOO, YOSHIHARU
; APPLICANT: YAMAGUCHI, KAZUO
; APPLICANT: YOSHIDA, Hajime
; APPLICANT: YOSHINORI, KOMATSU
; TITLE OF INVENTION: NOVEL POLYPEPTIDES
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,411
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 306799/86
; FILING DATE: 23-DEC-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 51357/88
; FILING DATE: 04-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 80088/88
; FILING DATE: 31-MAR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR
; REGISTRATION NUMBER: 25327
; REFERENCE/DOCKET NUMBER: 249-73
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100

```

```
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..522
; OTHER INFORMATION: /product= "HUMAN GRANULOCYTE COLONY
; OTHER INFORMATION: STIMULATING FACTOR"
US-08-434-411-1

Query Match 93.3%; Score 486.8; DB 1; Length 525;
Best Local Similarity 98.6%; Pred. No. 2.3e-109;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACCACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGGC 84
DB 25 CTGCCCCAGAGCTTCCTGCTCAAGTGTAGAGCAAGTGAGGAAGATCCAGGGCGATGGC 84
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
DB 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCACCCAGCGGCC 204
DB 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCACCCAGCGGCC 204
QY 205 CTGACGTGGCAGGCTCTTAAAGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGGC 264
DB 205 CTGACGTGGCAGGCTCTTAAAGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGGC 264
QY 265 CTGACGGCCCTGGAAGGATCTCCCGGAGTTGGGTCCTTGGACACTTGCACACTGCAGCTG 324
DB 265 CTGACGGCCCTGGAAGGATCTCCCGGAGTTGGGTCCTTGGACACTTGCACACTGCAGCTG 324
QY 325 GACGTGCGGAGCTTGGCACCACCATCTGCAGCAGATGGAAGACTGGAATGGCCCT 384
DB 325 GACGTGCGGAGCTTGGCACCACCATCTGCAGCAGATGGAAGACTGGAATGGCCCT 384
QY 385 GCCCTGACAGCCACCCAGGGTGCCATCGCGGCTTCGCCCTCTGCTTCCAGCGCCGGCA 444
DB 385 GCCCTGACAGCCACCCAGGGTGCCATCGCGGCTTCGCCCTCTGCTTCCAGCGCCGGCA 444
QY 445 GGAGGGTCTTAGTTGGCTCCCATCTCAGAGCTTCTGAGGAGTGTGTACCGCGTTCTA 504
DB 445 GGAGGGTCTTAGTTGGCTCCCATCTCAGAGCTTCTGAGGAGTGTGTACCGCGTTCTA 504
QY 505 CGCCACCTTGGCCAGGCC 522
DB 505 CGCCACCTTGGCCAGGCC 522
```

## RESULT 2

```
US-08-434-402-1
; Sequence 1, Application US/08434402
; Patent No. 5714581
```

## GENERAL INFORMATION:

```
; APPLICANT: KUGA, TETSURO
; APPLICANT: MIYAJI, HIROMASA
; APPLICANT: SATO, MORIYUKI
; APPLICANT: OKABE, MASAMI
; APPLICANT: MORIMOTO, MAKOTO
; APPLICANT: ITOH, SEIGA
; APPLICANT: YAMASAKI, MOTOO
; APPLICANT: YOKOO, YOSHIHARU
; APPLICANT: YAMAGUCHI, KAZUO
; APPLICANT: YOSHIDA, HAJIME
; APPLICANT: YOSHINORI, KOMATSU
```

```
; TITLE OF INVENTION: NOVEL POLYPEPTIDES
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,402
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 306799/86
; FILING DATE: 23-DEC-1986
; APPLICATION NUMBER: JP 51357/88
; FILING DATE: 04-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 80088/88
; FILING DATE: 31-MAR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR
; REGISTRATION NUMBER: 25327
; REFERENCE/DOCKET NUMBER: 249-72
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..522
; OTHER INFORMATION: /product= "HUMAN GRANULOCYTE COLONY
; OTHER INFORMATION: STIMULATING FACTOR"
US-08-434-402-1
```

```
Query Match 93.3%; Score 486.8; DB 1; Length 525;
Best Local Similarity 98.6%; Pred. No. 2.3e-109;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACCACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGGC 84
DB 25 CTGCCCCAGAGCTTCCTGCTCAAGTGTAGAGCAAGTGAGGAAGATCCAGGGCGATGGC 84
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
DB 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCCGAGGAGCTGGT 144
QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCACCCAGCGGCC 204
DB 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCACCCAGCGGCC 204
QY 205 CTGACGTGGCAGGCTCTTAAAGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGGC 264
DB 205 CTGACGTGGCAGGCTCTTAAAGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGGC 264
QY 265 CTGACGGCCCTGGAAGGATCTCCCGGAGTTGGGTCCTTGGACACTTGCACACTGCAGCTG 324
DB 265 CTGACGGCCCTGGAAGGATCTCCCGGAGTTGGGTCCTTGGACACTTGCACACTGCAGCTG 324
```

Qy	325	GAGTCGCCGAC	TTGCCACCACCATCTTGGCAGCAGATGGAAGAACTGGGAATGGCCCT	384
Db	325	GAGTCGCCGAC	TTTGGCACCAACCATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCT	384
Qy	385	GCCTTCGAGCC	ACCACGAGGTGCCATGCCGGCCTTGCCTCTGCTTTCAGCGCCGGCA	444
Db	385	GCCTTCGAGCC	ACCACGAGGTGCCATGCCGGCCTTGCCTCTGCTTTCAGCGCCGGCA	444
Qy	445	GGAGGGTCTCT	AGTTGGCTCCCATCTGCAGAGCTTCTCTGGAGGTGCTGACCGCTTCTA	504
Db	445	GGAGGGTCTCT	AGTTGGCTCCCATCTGCAGAGCTTCTCTGGAGGTGCTGACCGCTTCTA	504
Qy	505	CGCCACTTGC	CCAGCCC	522
Db	505	CGCCACTTGC	CCAGCCC	522

RESULT 3  
US-08-783-288-1  
; Sequence 1, Application US/08783288  
; Patent No. 5795968  
; GENERAL INFORMATION:  
; APPLICANT: KUGA, TETSURO  
; APPLICANT: MIYAJI, HIROMASA  
; APPLICANT: SATO, MORIYUKI  
; APPLICANT: OKABE, MASAMI  
; APPLICANT: MORIMOTO, MAKOTO  
; APPLICANT: ITOH, SEIGA  
; APPLICANT: YAMASAKI, MOTOO  
; APPLICANT: YOKOO, YOSHIMARU  
; APPLICANT: YAMAGUCHI, KAZUO  
; APPLICANT: YOSHIDA, HAJIME  
; APPLICANT: YOSHINORI, KOMATSU  
; TITLE OF INVENTION: NOVEL POLYPEPTIDES  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/783,288  
; FILING DATE: 10-JAN-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/434,411  
; FILING DATE: 03-MAY-1995  
; APPLICATION NUMBER: JP 306799/86  
; FILING DATE: 23-DEC-1986  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 51357/88  
; FILING DATE: 04-MAR-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 80088/88  
; FILING DATE: 31-MAR-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAWFORD, ARTHUR  
; REGISTRATION NUMBER: 25327  
; REFERENCE/DOCKET NUMBER: 249-73  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)816-4000  
; TELEFAX: (703)816-4100  
; TELEX: 200797 NIXN UR  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 525 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..522
; OTHER INFORMATION: /product= "HUMAN GRANULOCYTE COLONY
; OTHER INFORMATION: STIMULATING FACTOR"
; US-08-783-288-1

Query Match          93.3%; Score 486.8; DB 1; Length 525;
Best Local Similarity 98.6%; Pred. No. 2.3e-109;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACACAGAGCTTCCTTTTAAAAAGCTTAGAGCAAGTTAGAGAAAGATCCAGGGCGGATGGC 84
   |||
Db 25 CTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAAGTAGAGAAAGATCCAGGGCGGATGGC 84

QY 85 GCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAAGCTGTGCCACCCGAGGAGCTGGTG 144
   |||
Db 85 GCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAAGCTGTGCCACCCGAGGAGCTGGTG 144

QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCGACGAGGCC 204
   |||
Db 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCGACGAGGCC 204

QY 205 CTGAGCTGGCAGAGCTGCTTTGAGCAAACTCCATAGCGGCTTTTCTCTACCAAGGGCTC 264
   |||
Db 205 CTGAGCTGGCAGAGCTGCTTTGAGCAAACTCCATAGCGGCTTTTCTCTACCAAGGGCTC 264

QY 265 CTGAGCGCCCTGGAAGGAGTCTCCCGAGCTGGGTCGCCACCTGGACACACTGCAGCTG 324
   |||
Db 265 CTGAGCGCCCTGGAAGGAGTCTCCCGAGCTGGGTCGCCACCTGGACACACTGCAGCTG 324

QY 325 GAGCTGCCCGACTTTGGCCACCACCATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCT 384
   |||
Db 325 GAGCTGCCCGACTTTGGCCACCACCATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCT 384

QY 385 GCCCTGAGCCGCCAGGTGCCATCGCGGCTTCGCCTCTGCTTTCCAGCGCCGGGCA 444
   |||
Db 385 GCCCTGAGCCGCCAGGTGCCATCGCGGCTTCGCCTCTGCTTTCCAGCGCCGGGCA 444

QY 445 GGAGGGTCTCTAGTTGGCTCCCATCTCTGAGAGCTTCCTGGAGGTGTGCTACCGGTTCTA 504
   |||
Db 445 GGAGGGTCTCTAGTTGGCTCCCATCTCTGAGAGCTTCCTGGAGGTGTGCTACCGGTTCTA 504

QY 505 CGCCACCTTGGCCAGCCC 522
   |||
Db 505 CGCCACCTTGGCCAGCCC 522

RESULT 4
US-08-890-640-1
; Sequence 1, Application US/08890640
; Patent No. 5994518
; GENERAL INFORMATION:
; APPLICANT: KUGA, TETSURO
; APPLICANT: MIYAJI, HIROMASA
; APPLICANT: SATO, MORIYUKI
; APPLICANT: OKABE, MASAMI
; APPLICANT: MORIMOTO, MAKOTO
; APPLICANT: ITOH, SEIGA
; APPLICANT: YAMASAKI, MOTOO
; APPLICANT: YOKOO, YOSHIHARU
; APPLICANT: YAMAGUCHI, KAZUO
; APPLICANT: YOSHIDA, HAJIME
; APPLICANT: YOSHINORI, KOMATSU
; TITLE OF INVENTION: NOVEL POLYPEPTIDES
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:

```



Db 505 cgccacctgtgccagccc 522  
|||||

RESULT 6  
US-08-469-318-177  
; Sequence 177, Application US/08469318  
; Patent No. 6022535  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 196  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,318  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/446,872  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 177:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 546 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-469-318-177

Query Match 90.9%; Score 474.6; DB 3; Length 546;  
Best Local Similarity 94.4%; Pred. No. 2.le-106;  
Matches 492; Conservative 0; Mismatches 29; Indels 0; Gaps 0;  
QY 2 CGCCCAACATATCGCGCTCGAGCTTACCACAGAGCTTCTTTTAAAGCTTAGAGCAAG 61  
DB 8 CACCATTAGGCTTCCAGCTCCCTGCCAGAGCTTCTGCTCAAGTCTTAGAGCAAG 67  
QY 62 TGAGGAAGATCCAGGGCGATGCGCAGCTCCAGGAGAAGCTGTGTGCCACTTACAAGC 121  
DB 68 TGAGGAAGATCCAGGGCGATGCGCAGCTCCAGGAGAAGCTGTGTGCCACTTACAAGC 127  
QY 122 TGTGCCACCCCGAGAGCTGTGTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 181  
DB 128 TGTGCCACCCCGAGAGCTGTGTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 187  
QY 182 TGAGCAGCTGCCCGCAGGCGCTTGCAGCTGGCAGGCTTGTGAGCCAACTCCATAGCG 241  
DB 188 TGAGCTCTGCCCGCAGGCGCTTGCAGCTGGCAGGCTTGTGAGCCAACTCCATAGCG 247  
QY 242 GCCTTTTCTCTACAGGGGCTCTCTGAGGCGCTTGAAGGGATCTCCCGGAGTTGGGTC 301  
DB 248 GCCTTTTCTCTACAGGGGCTCTCTGAGGCGCTTGAAGGGATATCCCGGAGTTGGGTC 307  
QY 302 CCACCTTGACACACTGACGTGGACGTGCGGACTTTGCCACCACTCTGCGAGCAGA 361  
DB 308 CCACCTTGACACACTGACGTGGACGTGCGGACTTTGCCACCACTCTGCGAGCAGA 367  
QY 362 TGAAGAAGCTGGGAATGCCCTGCCCTGCAGCCACCCAGGCTGCCATGCCGGCTTCG 421  
DB 368 TGAAGAAGCTGGGAATGCCCTGCCCTGCAGCCACCCAGGCTGCCATGCCGGCTTCG 427  
QY 422 CCTCTGCTTTCAGCGCGGGCAGAGGGGTCTTAGTTGCTTCCCATCTGCGAGAGCTTCC 481  
DB 428 CCTCTGCTTTCAGCGCGGGCAGAGGGGTCTTAGTTGCTTCCCATCTGCGAGAGCTTCC 487  
QY 482 TGGAGGTGCTGACCGCGTTCTACGCCACCTTGGCCAGGCC 522  
|||||

Db 488 TGGAGGTGCTGACCGCGTTCTACGCCACCTTGGCCAGGCC 528

RESULT 7  
US-08-469-318-178  
; Sequence 178, Application US/08469318  
; Patent No. 6022535  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 196  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,318  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/446,872  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 178:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 546 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-469-318-178

Query Match 90.9%; Score 474.6; DB 3; Length 546;  
Best Local Similarity 94.4%; Pred. No. 2.le-106;  
Matches 492; Conservative 0; Mismatches 29; Indels 0; Gaps 0;  
QY 2 CGCCCAACATATCGCGCTCGAGCTTACCACAGAGCTTCTTTTAAAGCTTAGAGCAAG 61  
DB 8 CACCATTAGGCTTCCAGCTCCCTGCCAGAGCTTCTGCTCAAGTCTTAGAGCAAG 67  
QY 62 TGAGGAAGATCCAGGGCGATGCGCAGCTCCAGGAGAAGCTGTGTGCCACTTACAAGC 121  
DB 68 TGAGGAAGATCCAGGGCGATGCGCAGCTCCAGGAGAAGCTGTGTGCCACTTACAAGC 127  
QY 122 TGTGCCACCCCGAGAGCTGTGTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 181  
DB 128 TGTGCCACCCCGAGAGCTGTGTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 187  
QY 182 TGAGCAGCTGCCCGCAGGCGCTTGCAGCTGGCAGGCTTGTGAGCCAACTCCATAGCG 241  
DB 188 TGAGCTCTGCCCGCAGGCGCTTGCAGCTGGCAGGCTTGTGAGCCAACTCCATAGCG 247  
QY 242 GCCTTTTCTCTACAGGGGCTCTCTGAGGCGCTTGAAGGGATCTCCCGGAGTTGGGTC 301  
DB 248 GCCTTTTCTCTACAGGGGCTCTCTGAGGCGCTTGAAGGGATATCCCGGAGTTGGGTC 307  
QY 302 CCACCTTGACACACTGACGTGGACGTGCGGACTTTGCCACCACTCTGCGAGCAGA 361  
DB 308 CCACCTTGACACACTGACGTGGACGTGCGGACTTTGCCACCACTCTGCGAGCAGA 367  
QY 362 TGAAGAAGCTGGGAATGCCCTGCCCTGCAGCCACCCAGGCTGCCATGCCGGCTTCG 421  
DB 368 TGAAGAAGCTGGGAATGCCCTGCCCTGCAGCCACCCAGGCTGCCATGCCGGCTTCG 427  
QY 422 CCTCTGCTTTCAGCGCGGGCAGAGGGGTCTTAGTTGCTTCCCATCTGCGAGAGCTTCC 481  
DB 428 CCTCTGCTTTCAGCGCGGGCAGAGGGGTCTTAGTTGCTTCCCATCTGCGAGAGCTTCC 487  
QY 482 TGGAGGTGCTGACCGCGTTCTACGCCACCTTGGCCAGGCC 522  
|||||



TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-468-609A-178

Query Match 90.9%; Score 474.6; DB 3; Length 546;  
Best Local Similarity 94.4%; Pred. No. 2.1e-106;  
Matches 492; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 2 CGCCAAACATATCGCGCCTCGAGTCTTACCACAGAGCTTCTTTTAAAGCTTAGAGCAAG 61  
Db 8 CACCATTAGGACCTGCCAGTCCCTGCCCCAGAGCTTCTGCTCAAGTGTCTTAGAGCAAG 67  
Qy 62 TGAGGAAGATCCAGGGCGATGGCGAGGCTCCAGGAGAAGCTGTGTGCCACCTACAAGC 121  
Db 68 TGAGGAAGATCCAGGGCGATGGCGAGGCTCCAGGAGAAGCTGTGTGCCACCTACAAGC 127  
Qy 122 TGTGCCACCCCGAGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 181  
Db 128 TGTGCCACCCCGAGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 187  
Qy 182 TGAGCAGCTGCCCGCCAGCAGCCCTGACAGCTGGCAGGCTGTGAGCCAACTCCATAGCG 241  
Db 188 TGAGCTCTGCCCCAGCAGCCCTGACAGCTGGCAGGCTGTGAGCCAACTCCATAGCG 247  
Qy 242 GCCTTTTCTCTACAGGGGCTCTGACAGGCTTGGAGGGATCTCCCGGAGTTGGGTC 301  
Db 248 GCCTTTTCTCTACAGGGGCTCTGACAGGCTTGGAGGGATATCCCGGAGTTGGGTC 307  
Qy 302 CCACCTTGGACACACTGCAGCTGGACGTCGCGGACTTTGCCACCACTCTGGCAGCAGA 361  
Db 308 CCACCTTGGACACACTGCAGCTGGACGTCGCGGACTTTGCCACCACTCTGGCAGCAGA 367  
Qy 362 TGGAGAACTGGGAATGCCCTGCGCTGACGCCACCCAGGTCGCCATGCCGCTTCG 421  
Db 368 TGGAGAACTGGGAATGCCCTGCGCTGACGCCACCCAGGTCGCCATGCCGCTTCG 427  
Qy 422 CCTCTGCTTCCAGCGCGGGCAGGAGGGTCTAGTTGGCTTCCCATCTGCAGAGCTTCC 481  
Db 428 CCTCTGCTTCCAGCGCGGGCAGGAGGGTCTAGTTGGCTTCCCATCTGCAGAGCTTCC 487  
Qy 482 TGGAGGTGCTGACGGCTTCTACGCCACCTTGGCCAGCCC 522  
Db 488 TGGAGGTGCTGACGGCTTCTACGCCACCTTGGCCAGCCC 528

## RESULT 10

PCT-US95-01185-177

; Sequence 177, Application PC/TUS9501185

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion

; TITLE OF INVENTION: Protein

; NUMBER OF SEQUENCES: 196

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/01185

; FILING DATE: 02-FEB-1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/192325

; FILING DATE: 14-FEB-1994

; INFORMATION FOR SEQ ID NO: 177:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 546 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

PCT-US95-01185-177

Query Match 90.9%; Score 474.6; DB 5; Length 546;  
Best Local Similarity 94.4%; Pred. No. 2.1e-106;  
Matches 492; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 2 CGCCAAACATATCGCGCCTCGAGTCTTACCACAGAGCTTCTTTTAAAGCTTAGAGCAAG 61  
Db 8 CACCATTAGGACCTGCCAGTCCCTGCCCCAGAGCTTCTGCTCAAGTGTCTTAGAGCAAG 67  
Qy 62 TGAGGAAGATCCAGGGCGATGGCGAGGCTCCAGGAGAAGCTGTGTGCCACCTACAAGC 121  
Db 68 TGAGGAAGATCCAGGGCGATGGCGAGGCTCCAGGAGAAGCTGTGTGCCACCTACAAGC 127  
Qy 122 TGTGCCACCCCGAGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 181  
Db 128 TGTGCCACCCCGAGAGCTGGTGTCTCGGACACTCTCTGGGCATCCCTGGGCTCCCC 187  
Qy 182 TGAGCAGCTGCCCGCCAGCAGCCCTGACAGCTGGCAGGCTGTGAGCCAACTCCATAGCG 241  
Db 188 TGAGCTCTGCCCCAGCAGCCCTGACAGCTGGCAGGCTGTGAGCCAACTCCATAGCG 247  
Qy 242 GCCTTTTCTCTACAGGGGCTCTGACAGGCTTGGAGGGATCTCCCGGAGTTGGGTC 301  
Db 248 GCCTTTTCTCTACAGGGGCTCTGACAGGCTTGGAGGGATATCCCGGAGTTGGGTC 307  
Qy 302 CCACCTTGGACACACTGCAGCTGGACGTCGCGGACTTTGCCACCACTCTGGCAGCAGA 361  
Db 308 CCACCTTGGACACACTGCAGCTGGACGTCGCGGACTTTGCCACCACTCTGGCAGCAGA 367  
Qy 362 TGGAGAACTGGGAATGCCCTGCGCTGACGCCACCCAGGTCGCCATGCCGCTTCG 421  
Db 368 TGGAGAACTGGGAATGCCCTGCGCTGACGCCACCCAGGTCGCCATGCCGCTTCG 427  
Qy 422 CCTCTGCTTCCAGCGCGGGCAGGAGGGTCTAGTTGGCTTCCCATCTGCAGAGCTTCC 481  
Db 428 CCTCTGCTTCCAGCGCGGGCAGGAGGGTCTAGTTGGCTTCCCATCTGCAGAGCTTCC 487  
Qy 482 TGGAGGTGCTGACGGCTTCTACGCCACCTTGGCCAGCCC 522  
Db 488 TGGAGGTGCTGACGGCTTCTACGCCACCTTGGCCAGCCC 528

## RESULT 11

PCT-US95-01185-178

; Sequence 178, Application PC/TUS9501185

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion

; TITLE OF INVENTION: Protein

; NUMBER OF SEQUENCES: 196

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/01185

; FILING DATE: 02-FEB-1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/192325

; FILING DATE: 14-FEB-1994

; INFORMATION FOR SEQ ID NO: 178:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 546 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

PCT-US95-01185-178



Best Local Similarity 94.4%; Pred. No. 2.3e-106;		
Matches 492;	Conservative 0;	Mismatches 29; Indels 0; Gaps 0;
QY	2	CGCCCAACATATCGCGCCCTCGAGTCTTACCACAGAGCTTCCCTTTTAAAAAGCTTTAGAGCAAG 61
Db	8	CACCATAGGCCCTGSCAGACTCCCTTGCCCCAGAGCTTCTGCTCAAGTGTCTTAGAGCAAG 67
QY	62	TGAGGAAGATCATAGGCGGATGGCGAGCGCTCCAGGAGAAAGTGTGTGCCACCTACAAAGC 121
Db	68	TGAGGAAGATCATAGGCGGATGGCGAGCGCTCCAGGAGAAAGTGTGTGCCACCTACAAAGC 127
QY	122	TGTGCCACCCCGAGGAGCTGGTGTCTCGGACACTCTCTGGGATCCCTTGGGCTCCCC 181
Db	128	TGTGCCACCCCGAGGAGCTGGTGTCTCTCGGACACTCTCTGGGATCCCTTGGGCTCCCC 187
QY	182	TGAGCAGCTGGCCCCAGCCAGCGCCCTGCGAGCTTGGCAGGCTGCTTGAGCCAACTCCATAGCG 241
Db	188	TGAGCTCTCTGCCAGCCAGCGCCCTGCGAGCTTGGCAGGCTGCTTGAGCCAACTCCATAGCG 247
QY	242	GCCCTTTTCTCTACGAGGGTCTCTGCGAGGCCCTGGGAAGGATCTCCCGCCAGTTGGGTC 301
Db	248	GCCCTTTTCTCTACGAGGGTCTCTGCGAGGCCCTGGGAAGGATATCCCGCCAGTTGGGTC 307
QY	302	CCACCTTGGACACACTGCGAGCTGGAGGTCGCCGACTTTGCCACCACCACTCTGGCAGCAGA 361
Db	308	CCACCTTGGACACACTGCGAGCTGGAGGTCGCCGACTTTGCCACCACCACTCTGGCAGCAGA 367
QY	362	TGGAAGAACTGGGAATGGCCCCCTGCCCTGCGAGCCCCACCCAGGGTGCCATGCGCGCCTTCG 421
Db	368	TGGAAGAACTGGGAATGGCCCCCTGCCCTGCGAGCCCCACCCAGGGTGCCATGCGCGCCTTCG 427
QY	422	CCTCTGCTTTCCAGGGCGGGCAGAGGGGTCTAGTTGGCTCCCATCTGCAGAGCTTCC 481
Db	428	CCTCTGCTTTCCAGGGCGGGCAGAGGGGTCTCTGGTTGTAGCCATCTGCAGAGCTTCC 487
QY	482	TGGAGGTCTCTACCGCGCTTCTACGCCACCTTGGCCAGCCCC 522
Db	488	TGGAGGTCTCTACCGCGCTTCTACGCCACCTTGGCCAGCCCC 528
RESULT 14		
US-08-469-318-78		
; Sequence 78, Application US/08469318		
; Patent No. 6022535		
; GENERAL INFORMATION:		
; APPLICANT:		
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion		
; TITLE OF INVENTION: Protein		
; NUMBER OF SEQUENCES: 196		
; COMPUTER READABLE FORM:		
; MEDIUM TYPE: Floppy disk		
; COMPUTER: IBM PC compatible		
; OPERATING SYSTEM: PC-DOS/MS-DOS		
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)		
; CURRENT APPLICATION DATA:		
; APPLICATION NUMBER: US/08/469,318		
; FILING DATE:		
; CLASSIFICATION:		
; PRIOR APPLICATION DATA:		
; APPLICATION NUMBER: 08/446,872		
; FILING DATE:		
; INFORMATION FOR SEQ ID NO: 78:		
; SEQUENCE CHARACTERISTICS:		
; LENGTH: 921 base pairs		
; TYPE: nucleic acid		
; STRANDEDNESS: double		
; TOPOLOGY: linear		
; MOLECULE TYPE: DNA (genomic)		
US-08-469-318-78		

Query Match	90.98;	Score 474.6;	DB 3;	Length 921;
Best Local Similarity	94.48;	Pred. No. 2.3e-106;		

Matches	492;	Conservative	0;	Mismatches	29;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	-----	--------	----	------	----

  

QY	2	CGCCAAACATATCGCGCCCTCGAGTCTACACAGAGAGTCTCTTTTAAAAAGCTTAGAGCAAG	61
Db	8		
QY	62	TCAGAAAGATCAGGGCGATGCGCAGGCTCCAGGAGAAAGCTGTGTGCCACCTTACAAGC	121
Db	68	TCAGAAAGATCAGGGCGATGCGCAGGCTCCAGGAGAAAGCTGTGTGCCACCTTACAAGC	127
QY	122	TGTGGCACCCCGAGGAGCTGTGTGCTGCTCGGACACTCTCTGGGATCCCTTGGGCTCCCC	181
Db	128	TGTGCCACCCCGAGGAGCTGTGTGCTGCTCGGACACTCTCTGGGATCCCTTGGGCTCCCC	187
QY	182	TGAGCAGTGTGCCCGACGAGCCCTGTCAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCG	241
Db	188	TGAGCTCTGTGCCCGACGAGCCCTGTCAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCG	247
QY	242	GCCTTTTCTCTACACAGGGCTCTGCAAGGCCCTTGGAGGGATATCCCGCAGTTGGGTC	301
Db	248	GCCTTTTCTCTACACAGGGCTCTGCAAGGCCCTTGGAGGGATATCCCGCAGTTGGGTC	307
QY	302	CCACCTTGGACACACTGTCAGCTGGACGCTGCGCGACTTTGCCACCACTCTGGCAGCAGA	361
Db	308	CCACCTTGGACACACTGTCAGCTGGACGCTGCGCGACTTTGCCACCACTCTGGCAGCAGA	367
QY	362	TGGAAGAACTGGGAATGCCCTTCCCTGTCAGGCCACCCAGGCTGCCATGCCGCCCTTCG	421
Db	368	TGGAAGAACTGGGAATGCCCTTCCCTGTCAGGCCACCCAGGCTGCCATGCCGCCCTTCG	427
QY	422	CCTCTGCTTTCCAGCGCGCGGAGGAGGGTCTAGTTGCCCTCCCATCTGCAGAGCTTCC	481
Db	428	CCTCTGCTTTCCAGCGCGGAGGAGGGTCTAGTTGCCCTCCCATCTGCAGAGCTTCC	487
QY	482	TGGAGGTGCTGTACCGGCTTCTACGCCACCTTGCCCGAGCCC	522
Db	488	TGGAGGTGCTGTACCGGCTTCTACGCCACCTTGCCCGAGCCC	528

  

RESULT	15
US-08-469-318-84	
Sequence 84, Application US/08469318	
Patent No. 6022535	
GENERAL INFORMATION:	
APPLICANT:	
TITLE OF INVENTION:	Multivariant IL-3 Hematopoiesis Fusion
TYPE OF INVENTION:	Protein
NUMBER OF SEQUENCES:	196
COMPUTER READABLE FORM:	
MEDIUM TYPE:	Floppy disk
COMPUTER:	IBM PC compatible
OPERATING SYSTEM:	PC-DOS/MS-DOS
SOFTWARE:	PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:	
APPLICATION NUMBER:	US/08/469,318
FILING DATE:	
CLASSIFICATION:	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER:	08/446,872
FILING DATE:	
INFORMATION FOR SEQ ID NO: 84:	
SEQUENCE CHARACTERISTICS:	
LENGTH:	921 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	double
TOPOLOGY:	linear
MOLECULE TYPE:	DNA (genomic)
US-08-469-318-84	

Query Match	90.9%;	Score 474.6;	DB 3;	Length 921;
Best Local Similarity	94.4%;	Pred. NO. 2.3e-106;		
Matches 492; Conservative	0;	Mismatches 29;	Indels 0;	Gaps 0;



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2002, 08:59:09 ; Search time 262.13 seconds  
(without alignments)  
3419.026 Million cell updates/sec

Title: US-09-680-514-4\_COPY\_526\_1047  
Perfect score: 522  
Sequence: 1 GCGCCAAATATCGCGCTC.....TACGCCACCTTGCCAGCCC 522

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	522	100.0	1047	17	AAT41786 Fusion peptide #1
2	522	100.0	1095	17	AAT41788 Fusion peptide #3
3	520.4	99.7	1083	17	AAT41787 Fusion peptide #2
4	514	98.5	525	11	AA004482 Plasmid PASN6 enco
5	510.8	97.9	525	11	AA004484 Plasmid PASN145 en
6	506	96.9	525	11	AA004481 Plasmid PAS28 enco
7	486.8	93.3	525	9	AAN80947 G-CSF gene isolate
8	486.8	93.3	531	22	AAD19772 Human hc-CSF gene
9	486.8	93.3	615	22	AAD19771 Human hc-CSF gene

10	486.8	93.3	1520	8	AAN70223 Plasmid pBRV2 inse
11	486.8	93.3	1520	8	AAN71320 Sequence encoding
12	486.8	93.3	1521	7	Plasmid pBRV2 inse
13	486.8	93.3	1525	9	Sequence encoding
14	486.8	93.3	1525	10	Plasmid pP12 contg
15	485.2	93.0	1415	8	Sequence of human
16	484.2	92.8	644	22	AAI71848 Recombinant human
17	474.6	90.9	546	16	AAQ97206 pMONI3499 DNA sequ
18	474.6	90.9	546	16	AAQ97205 pMONI3010 DNA sequ
19	474.6	90.9	546	21	AAQ37772 Human G-CSF mutant
20	474.6	90.9	546	21	AAQ37773 pMONI3026 DNA enco
21	474.6	90.9	921	16	AAQ97186 pMONI3043 DNA enco
22	474.6	90.9	921	16	AAQ97189 pMONI3151 DNA enco
23	474.6	90.9	921	16	AAQ97192 pMONI3066 DNA enco
24	474.6	90.9	921	16	AAQ97198 Human interleukin-
25	474.6	90.9	921	21	AAQ03740 Human interleukin-
26	474.6	90.9	921	21	AAQ03743 Human interleukin-
27	474.6	90.9	921	21	AAQ03746 Human interleukin-
28	474.6	90.9	921	21	AAQ03752 Human interleukin-
29	474.6	90.9	923	20	AA238850 Human granulocytic
30	474.6	90.9	966	16	AAQ97184 pMONI3058 DNA enco
31	474.6	90.9	966	16	AAQ97185 pMONI3060 DNA enco
32	474.6	90.9	966	16	AAQ97187 pMONI3063 DNA enco
33	474.6	90.9	966	16	AAQ97191 pMONI3045 DNA enco
34	474.6	90.9	966	16	AAQ97193 pMONI3152 DNA enco
35	474.6	90.9	966	16	AAQ97182 pMONI3034 DNA enco
36	474.6	90.9	966	21	AAQ03736 Human interleukin-
37	474.6	90.9	966	21	AAQ03738 Human interleukin-
38	474.6	90.9	966	21	AAQ03739 Human interleukin-
39	474.6	90.9	966	21	AAQ03741 Human interleukin-
40	474.6	90.9	966	21	AAQ03745 Human interleukin-
41	474.6	90.9	966	21	AAQ03747 Human interleukin-
42	474.6	90.9	1017	16	AAQ97181 pMONI5937 DNA enco
43	474.6	90.9	1017	21	AAQ03735 Human interleukin-
44	474.6	90.9	1047	16	AAQ97188 pMONI3084 DNA enco
45	474.6	90.9	1047	16	AAQ97190 pMONI3044 DNA enco

ALIGNMENTS

RESULT 1  
AAT41786  
ID AAT41786 standard; DNA; 1047 BP.

XX AC AAT41786;

XX DT 01-JUL-1997 (first entry)

XX DE Fusion peptide #1 having G-CSF and TPO activity.

XX DE Fusion protein; human granulocyte colony stimulating factor; hc-CSF;

KW KW thrombopoietin; TPO; spacer peptide; blood platelet production;

KW leukocyte production; anaemia; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT sig\_peptide 1..63

FT mat\_peptide /\*tag= a

FT /\*tag= b

XX PN WO9634016-A1.

XX PD 31-OCT-1996.

XX PF 26-APR-1996; 96WO-JP01157.

XX PR 26-APR-1995; 95JP-0102625.

XX PA (KYOW ) KYOWA HAKKO KOGYO KK.

XX XX

PI Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;  
XX Uchida K, Yamasaki M, Yamashita K, Yokoi H;  
DR WPI: 1996-497573/49.  
XX P-PSDB; AAW00377.  
PT Fusion peptide having G-CSF and with thrombopoietin activity -  
PT optionally chemically modified with a poly:alkylene glycol, used for  
XX treatment of anaemia  
PS Claim 3; Page 46-48; 71pp; Japanese.  
XX The sequences given in AAT41786-88 encode fusion proteins which are  
CC composed of a peptide having human granulocyte colony stimulating  
CC factor (hG-CSF) activity fused with a peptide having thrombopoietin  
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from  
CC these by deletion, insertion or substitution of one or more amino  
CC acid residues are included within the scope of the invention. The  
CC fusion peptides stimulate blood platelet and leukocyte production  
CC and are useful in the treatment of anaemia.  
XX Sequence 1047 BP; 192 A; 351 C; 288 G; 216 T; 0 other;  
SQ  
Query Match 100.0%; Score 522; DB 17; Length 1047;  
Best Local Similarity 100.0%; Pred. No. 2.7e-109; Indels 0; Gaps 0;  
Matches 522; Conservative 0; Mismatches 0;  
QY 1 GCGCCAACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTTAAAGCTTAGACAA 60  
Db |||||||  
526 ggcgaacatcgctcgagtcagctaccacagagcttccttttaaaagcttagacaa 585  
QY 61 GTGAGGAGATCCAGGGGATGGCGAGCGCTCAGAGAGCTGTGTCACCTACAG 120  
Db |||||||  
586 gtgaggaagatccaggcgatggcgagcgctccagagagctgtgtgccactacaag 645  
QY 121 CTGTGCCACCCGAGGAGCTGCTGTCTGGACACTCTCTGGSCATCCCTGGGCTCCC 180  
Db |||||||  
646 ctgttccaccccgaggagctgtgtgtctgctcggacactctctg99catccccctggctccc 705  
QY 181 CTGAGCAGCTGCCCGCAGCGCCCTCGAGTGCGAGGCTGCTTGAGCCAACTCCATAGC 240  
Db |||||||  
706 ctgagcagctgcccagcagcgccctgcagctggcagctgtgtgagccaaactccatagc 765  
QY 241 GGCCTTTTCTCTACAGGGGCTCTCGAGGCCCTGGAAGGATCTCCCCGAGTTGGGT 300  
Db |||||||  
766 ggccttttctctaccaggggctcctgcagggcctggaaggagatctcccccgagttgggt 825  
QY 301 CCCACTTTGGACACTGCAGCTGGAGCTGCCGACTTTTGGCCACCATCTGGCAGCAG 360  
Db |||||||  
826 cccacttggacacactgcagctgagctgagctgcgcagcttttgcacacacatctggcagcag 885  
QY 361 ATGGAAGAACTGGGAATGGCCCTTGCCTCGAGCCACCCAGGCTGCCATGCCGCTTC 420  
Db |||||||  
886 atggaagaactgggaatggcccttgcctgcagccacccaggggtgccatgccgcttc 945  
QY 421 GCGTCTGTTTCCAGCGCGGGAGGAGGGTCTGTAGTTCCTCCATCTGCGAGCTTC 480  
Db |||||||  
946 gccctgtcttccagcgccggcaggagggtgctagttgtctcccatctgccagcttc 1005  
QY 481 CTGAGGTGTGCTACCGCTTCTACGCCACCTTTGCCAGGCC 522  
Db |||||||  
1006 ctggaggtgtgtaccgcttctacgcacacttggccagccc 1047  
RESULT 2  
ID AAT41788  
XX AAT41788 standard; DNA: 1095 BP.  
AC AAT41788;  
XX  
DT 01-JUL-1997 (first entry)  
XX

DE Fusion peptide #3 having G-CSF and TPO activity.  
XX  
KW Fusion protein; human granulocyte colony stimulating factor; hG-CSF;  
KW thrombopoietin; TPO; spacer peptide; blood platelet production;  
KW leukocyte production; anaemia; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT sig\_peptide 1..63 /\*tag= a  
FT mat\_peptide 64..1095 /\*tag= b  
FT  
FT  
XX W09634016-A1.  
XX  
XX 31-OCT-1996.  
XX  
XX 26-APR-1996; 96WO-JP011157.  
XX  
XX 26-APR-1995; 95JP-0102625.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
XX Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;  
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;  
XX  
XX WPI: 1996-497573/49.  
DR P-PSDB; AAW00379.  
XX  
XX Fusion peptide having G-CSF and with thrombopoietin activity -  
XX optionally chemically modified with a poly:alkylene glycol, used for  
XX treatment of anaemia  
XX  
XX Claim 3; Page 52-54; 71pp; Japanese.  
XX The sequences given in AAT41786-88 encode fusion proteins which are  
CC composed of a peptide having human granulocyte colony stimulating  
CC factor (hG-CSF) activity fused with a peptide having thrombopoietin  
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from  
CC these by deletion, insertion or substitution of one or more amino  
CC acid residues are included within the scope of the invention. The  
CC fusion peptides stimulate blood platelet and leukocyte production  
CC and are useful in the treatment of anaemia.  
XX  
XX Sequence 1095 BP; 194 A; 362 C; 311 G; 228 T; 0 other;  
SQ  
Query Match 100.0%; Score 522; DB 17; Length 1095;  
Best Local Similarity 100.0%; Pred. No. 2.7e-109; Indels 0; Gaps 0;  
Matches 522; Conservative 0; Mismatches 0;  
QY 1 GCGCCAACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTTAAAGCTTAGACAA 60  
Db |||||||  
574 ggcgaacatcgctcgagtcagctaccacagagcttccttttaaaagcttagacaa 633  
QY 61 GTGAGGAGATCCAGGGGATGGCGAGCGCTCAGAGAGCTGTGTCACCTACAG 120  
Db |||||||  
634 gtgaggaagatccaggcgatggcgagcgctccagagagctgtgtgccactacaag 693  
QY 121 CTGTGCCACCCGAGGAGCTGCTGTCTGGACACTCTCTGGSCATCCCTGGGCTCCC 180  
Db |||||||  
694 ctgtgccaccccgaggagctgtgtgtctcggacactctctg99catccccctggctccc 753  
QY 181 CTGAGCAGCTGCCCGCAGCGCCCTGCAGCTGGAGGCTGTGAGCCAACTCCATAGC 240  
Db |||||||  
754 ctgagcagctgcccgagccagcgccctgcagctggcagctgtgtgagccaaactccatagc 813  
QY 241 GGCCTTTTCTCTACCGAGGGCTCTCGAGGCCCTGGAAGGATCTCCCCGAGTTGGGT 300  
Db |||||||  
814 ggccttttctctaccaggggctcctgcagggcctggaaggagatctcccccgagttgggt 873  
QY 301 CCCACTTTGGACACTGCAGCTGGAGCTGCCGACTTTTGGCCACCATCTGGCAGCAG 360

|||||  
Db 874 cccaccttggacacactgcagctgagctgcgcgactttgccaccacatctggcagcag 933  
Qy 361 ATGGAAGAACTGGGAATGCGCCCTGCGCTGCAGGCCACCCAGGGTGCCATGCCGCGCTTC 420  
Db 934 atggaagaac-1999aatgcccccttgcctgcagccacccagggtgccatgcgcgcccttc 993  
Qy 421 GCCTCTGCTTTCCAGCGCGGAGGAGGGGTCTCTAGTTGGCTTCCCATCTGCAGAGCTTC 480  
Db 994 gccctgctttccagcgccggcagagggtcctagtgtgctcccatctgcagagcttc 1053  
Qy 481 CTGAGAGGTGCTACCGCGTCTTACGCCACCTTGCCCGAGCCC 522  
Db 1054 ctggagggtgtcgtaccgcgttctacgccacacttgccacgcc 1095

RESULT 3  
AA041787  
ID AA041787 standard; DNA; 1083 BP.

XX AA041787;  
XX  
DT 01-JUL-1997 (first entry)  
XX  
DE Fusion peptide #2 having G-CSF and TPO activity.  
XX  
KW Fusion protein; human granulocyte colony stimulating factor; hg-CSF;  
KW thrombopoietin; TPO; spacer peptide; blood platelet production;  
KW leukocyte production; anaemia; ds.  
XX  
OS Homo sapiens.

XX Key Location/Qualifiers  
FH sig\_peptide 1..63  
FT /\*tag= a  
FT mat\_peptide 64..1083  
FT /\*tag= b

XX W09634016-A1.  
XX  
XX 31-OCT-1996.  
XX  
XX 26-APR-1996; 96WO-JP01157.  
XX  
XX 26-APR-1995; 95JP-0102625.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.

XX Anazawa H, Konishi N, Shiotsu Y, Tamaoki T, Terasaki Y;  
PI Uchida K, Yamasaki M, Yamashita K, Yokoi H;  
XX  
XX WPI; 1996-497573/49.  
DR P-PSDB; AAW00378.

XX Fusion peptide having G-CSF and with thrombopoietin activity -  
PT optionally chemically modified with a poly(alkylene glycol, used for  
PT treatment of anaemia

XX Claim 3; Page 49-51; 71pp; Japanese.

XX The sequences given in AA041786-88 encode fusion proteins which are  
CC composed of a peptide having human granulocyte colony stimulating  
CC factor (hg-CSF) activity fused with a peptide having thrombopoietin  
CC (TPO) activity, opt. via a spacer peptide. Peptides derived from  
CC these by deletion, insertion or substitution of one or more amino  
CC acid residues are included within the scope of the invention. The  
CC fusion peptides stimulate blood platelet and leukocyte production  
CC and are useful in the treatment of anaemia.

XX Sequence 1083 BP; 197 A; 355 C; 305 G; 226 T; 0 other;

Query Match 99.7%; Score 520.4; DB 17; Length 1083;

Best Local Similarity 99.8%; Pred. No. 6.3e-109;  
Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GCGCCAAATATCGCGCCTCGAGTCTACACAGAGCTTCCCTTTTAAAAAGCTTAGAGCAA 60  
Db 562 gcaccaacatcgcgcctcagtgctacacagagctccttttaaaagcttagagcaa 621  
Qy 61 GTGAGGAAGATCCAGGGGATGGCGCAGTCTCCAGGAGAAGCTGTGTGCCACCTTACAAG 120  
Db 622 gtgaagaagatccaggcgatgagcgagcgtcccaaggagaagctgtgccaacctacaag 681  
Qy 121 CTGTGCCACCCCGAGAGAGCTGTGTCTGCTCGGACACTCTCTTGGGCATCCCTTGGGCTCC 180  
Db 682 ctgtgccaccccgaggagtggtgctgctcgacactctctgggcatccctcgggctccc 741  
Qy 181 CTGAGCAGCTGCCCGCAGCCAGCCCTGCAGCTGGCAGGCTGCTTGCAGCCAACTCCATAGC 240  
Db 742 ctgagcagctgccccagccagccctgcagctggcaggtgctgagccaaactccatagc 801  
Qy 241 GGCTTTTCTCTACAGGGGCTCCTGCAGGCCCTTGGAGGGATCTCCCCGAGTTGGGT 300  
Db 802 ggcttttctctaccagggtcctcctgcagccctggaaggatctccccgagttgggt 861  
Qy 301 CCCACCTTGGACACTGCAGCTGCAGCTGCGCGACTTTGGCACCACTCTGGCAGCAG 360  
Db 862 cccaccttggacacactgcagctggcagctgcgcgactttgccaccaccatctggcagcag 921  
Qy 361 ATGGAAGAACTGGGAATGSCCTCTGCAGGCCACCCAGGGTGCCATGCCGCGCTTC 420  
Db 922 atggaagaactggaatgccccctgcctgcagccacccacccagggcgccatgcggccctc 981  
Qy 421 GCCTCTGCTTTCCAGCGCGGAGGGGTCTCTAGTTGGCTTCCCATCTGCAGAGCTTC 480  
Db 982 gccctgctttccagcgccggcagagggttcctagttgctcccatctgcagagcttc 1041  
Qy 481 CTGAGAGGTGCTACCGCGTCTTACGCCACCTTGCCCGAGCCC 522  
Db 1042 ctggagggtgtcgtaccgcgttctacgccacacttgccacgcc 1083

RESULT 4  
AA04482  
ID AA04482 standard; DNA; 525 BP.  
XX  
XX AC AA04482;  
XX  
XX DT 04-OCT-1990 (first entry)  
XX  
XX DE Plasmid pASN6 encoding hg-CSF[ND28N6].  
XX KW Granulocyte stimulating factor; O-linked glycosylation; ss.  
XX OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH misc\_difference 1..3  
FT /\*tag= a  
FT /label=synthetic\_mutation  
FT /note="old seq (acc)"  
FT misc\_difference 7..18  
FT /\*tag= b  
FT /label=synthetic\_mutation  
FT /note="old seq (ctgggcccctgcc)"  
FT misc\_difference 49..51  
FT /\*tag= c  
FT /label=synthetic\_mutation  
FT /note="old seq (tgc)"  
XX  
XX EP370205-A.  
XX  
XX 30-MAY-1990.  
XX  
XX 28-SEP-1989; 89EP-0117981.

```
XX 29-SEP-1989; 89JP-0245705.
XX (KYOW ) KYOWA HAKKA KK.
XX PI Sasaki K, Nishi T, Yasumuru S, Sato M, Itoh S;
XX WPI; 1990-165029/22.
XX P-PSDB; AAR05114.
XX Polypeptide(s) with added carbohydrate chains - formed by
PT modification of amino acid sequence, used to improve
PT physiochemical properties and/or activities.
XX PS Disclosure; ; 30pp; English.
XX The sequence encodes a deriv. of mature hG-CSF, designated
CC hG-CSF[ND28N] which has 6 amino acid substns. This results in
CC one extra O-linked glycosylation site and one new N-linked
CC glycosylation site giving the new protein improved stability and
CC activity.
CC See also AAQ04481-86.
XX SQ Sequence 525 BP; 94 A; 183 C; 148 G; 100 T; 0 other;

Query Match 98.5%; Score 514; DB 11; Length 525;
Best Local Similarity 99.0%; Pred. No. 1.6e-107;
Matches 517; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCCACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTTAAAGCTTAGAGCAA 60
DB 1 gcaactacgtatcgcaactcagctgctaccacagagcttctctttaaagaagcttagagcaa 60

QY 61 CTGAGGAGATCCAGGCGGATGGCGAGCTCCAGGAGAGCTGTGCCACCTACAG 120
DB 61 gtgaggaaatcccaaggcgatggcgagcgctccagagagagctgtgtgccacctacaag 120

QY 121 CTGTGCCACCCGAGGAGCTGGTCTCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
DB 121 ctgtgccacccgagagagctgggtgctctcgagacctctctgggcatccctgggctccc 180

QY 181 CTGAGCAGCTGCCCGACGAGCGCTCGAGCTGCGAGGCTGTGAGCCAACTCCATAGC 240
DB 181 ctgagcagctgcccgacgagcgctcgagctgagagctgtgagccaaactccatagc 240

QY 241 GGCCTTTTCTTACAGGGGCTCTCGAGGCTCGAGGCTGAGGAGTCTCCCGAGTTGGT 300
DB 241 ggccttttcttaccaggggctctcgagggctcgaggggagatctccccccgagttgggt 300

QY 301 CCCACCTGGACACACTGCAGCTGGAGCTGCGCGACTTTTGCCACCACTCTGGCAGCAG 360
DB 301 cccacctggacacactgcagctggagctgagctgagcttggcaccaccttggcagcag 360

QY 361 ATGGAAGAACTGGGAATGGCCCTGCGCTCGAGCCACCCAGGCTGCGATGCGCGCTTC 420
DB 361 atggaagaactgggaatggccctgagccctgagccacccaggggtgcatgcccgccttc 420

QY 421 GCCTCTGCTTCCAGCGCGCGGAGGAGGCTCTTAGTTCCTCCATCTGCGAGAGCTTC 480
DB 421 gcctctgcttccagcgccgagggaggggtctctagttcctccatctgacagagcttc 480

QY 481 CTGAGGTGTGTACCGCTTCTACGCCACCTTGGCCAGGCC 522
DB 481 ctgagggtgtgtaccgcttctacgcccaccttgccagccc 522

RESULT 5
AAQ04484
ID AAQ04484 standard; DNA; 525 BP.
XX AAQ04484;
XX AAQ04484;
```

```
DT 04-OCT-1990 (first entry)
XX Plasmid pASN145 encoding hG-CSF[ND28N145].
XX KW Granulocyte stimulating factor; glycosylation; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT misc_difference 1..3 /*tag= a
FT /label=synthetic_mutation
FT /note="old seq (acc)"
FT misc_difference 7..18 /*tag= b
FT /label=synthetic_mutation
FT /note="old seq (ctgggcccctgccc)"
FT misc_difference 49..51 /*tag= c
FT /label=synthetic_mutation
FT /note="old seq (tgc)"
FT misc_difference 433..435 /*tag= d
FT /label=synthetic_mutation
FT /note="old seq (cag)"
FT misc_difference 439..441 /*tag= e
FT /label=synthetic_mutation
FT /note="old seq (cgg)"
XX EP370205-A.
XX 30-MAY-1990.
XX 28-SEP-1989; 89EP-0117981.
XX 29-SEP-1989; 89JP-0245705.
XX (KYOW ) KYOWA HAKKA KK.
XX Sasaki K, Nishi T, Yasumuru S, Sato M, Itoh S;
XX WPI; 1990-165029/22.
XX P-PSDB; AAR05115.
XX Polypeptide(s) with added carbohydrate chains - formed by
PT modification of amino acid sequence, used to improve
PT physiochemical properties and/or activities.
XX PS Disclosure; ; 0pp; English.
XX The sequence encodes a deriv. of mature hG-CSF, designated
CC hG-CSF[ND28N145] which has 7 amino acid substns. This results in
CC one extra O-linked glycosylation site and two new N-linked
CC glycosylation site giving the new protein improved stability and
CC activity.
CC See also AAQ04481-86.
XX SQ Sequence 525 BP; 96 A; 183 C; 145 G; 101 T; 0 other;

Query Match 97.9%; Score 510.8; DB 11; Length 525;
Best Local Similarity 98.7%; Pred. No. 8.6e-107;
Matches 515; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GCGCCACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTTAAAGCTTAGAGCAA 60
DB 1 gcaaccaacatatcgcgctcgagctctaccacagagcttctctttaaagaagcttagagcaa 60

QY 61 GTGAGGAGAGATCCAGGCGGATGGCGAGGCTCCAGGAGAGCTGTGTGCCACCTACAG 120
DB 61 gtgaggaaatcccaaggcgatggcgagcgctccagagagagctgtgtgccacctacaag 120
```

```
QY 121 CTGTGCCACCCGAGGAGCTGTGTGCTGCTGACACTCTCTGGGCATCCCTGGGCTCCC 180
Db 121 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 181 CTGACGAGTCCCGCAGCCGCTGACCTGGAGCTGGAGGCTGTTGAGCAACTCCATAGC 240
Db 181 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 241 GGCCTTTTCTCTACACAGGGCTCTGTGACGGCCCTGGAAGGAGATCTCCCGGAGTTGGGT 300
Db 241 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 301 CCCACTTGGACACACTGACGTGACGTGCGCCGACTTTGCCACCACTATCTGGCAGCAG 360
Db 301 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 361 ATGGAGAACTGGGAATGCCCTGCTCCCTGCGACGCCACCCAGGCTGCCATGCCGCTTC 420
Db 361 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 421 GCCTCTGCTTCTCAGCGCGGCGAGGAGGGTCTCTAGTTGCTTCCCTATCTGCAGAGCTTC 480
Db 421 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 481 CTGGAGGTGTGACCGGCTTCTACGCCACCTTGCCACCTTGCCCGAGCCC 522
Db 481 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

## RESULT 6

```
ID AA004481 standard; DNA; 525 BP.
XX AC AA004481;
XX 04-OCT-1990 (first entry)
XX Plasmid pAS28 encoding hG-CSF[ND28].
XX Granulocyte stimulating factor; O-linked glycosylation; ss.
XX Homo sapiens.
```

```
Key Location/Qualifiers
FT misc_difference 1..3 /*tag= a
FT /label=synthetic_mutation
FT /note="old seq (acc)"
FT misc_difference 7..15 /*tag= b
FT /label=synthetic_mutation
FT /note="old seq (ctgggacct)"
FT misc_difference 49..51 /*tag= c
FT /label=synthetic_mutation
FT /note="old seq (tgc)"
```

```
PN EP370205-A.
XX 30-MAY-1990.
XX 28-SEP-1989; 89EP-0117981.
XX 29-SEP-1989; 89JP-0245705.
XX (KYOW ) KYOWA HAKKA KK.
XX Sasaki K, Nishi T, Yasumuru S, Sato M, Itoh S;
XX WPI: 1990-165029/22.
XX P-PSDB; AAR05113.
XX Polypeptide(s) with added carbohydrate chains - formed by
```

```
PT modification of amino acid sequence, used to improve
XX physiochemical properties and/or activities.
PS Disclosure; ; 30pp; English.
XX The sequence encodes a deriv. of mature hG-CSF, designated
CC hG-CSF[N28] which has 5 amino acid substns. This results in
CC one extra O-linked glycosylation site giving the new protein
CC improved stability and activity.
XX See also AA004482-86.
SQ Sequence 525 BP; 94 A; 184 C; 149 G; 98 T; 0 other;
```

```
Query Match 96.9%; Score 506; DB 11; Length 525;
Best Local Similarity 98.1%; Pred. No. 1.le-105;
Matches 512; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GCGCCAAACATATCGCGCCTCGAGTCTACACAGAGCTCTCTTTTAAAGCTTAGAGCAA 60
Db 1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 61 GTGAGGAAGATCCAGGGCGATGGCGAGCGCTCCAGGAGAAAGCTGTGCCACCTACAG 120
Db 61 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 121 CTGTGCCACCCGAGGAGCTGTGTGCTGCTGACACTCTCTGGGCATCCCTGGGCTCCC 180
Db 121 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 181 CTGACGAGTCCCGCAGCCCTGCTGACCTGGAGGCTGTTGAGCAACTCCATAGC 240
Db 181 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 241 GGCCTTTTCTCTACACAGGGCTCTGTGACGGCCCTGGAAGGAGATCTCCCGGAGTTGGGT 300
Db 241 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 301 CCCACTTGGACACACTGACGTGACGTGCGCCGACTTTGCCACCACTATCTGGCAGCAG 360
Db 301 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 361 ATGGAGAACTGGGAATGCCCTGCTCCCTGCGACGCCACCCAGGCTGCCATGCCGCTTC 420
Db 361 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 421 GCCTCTGCTTCTCAGCGCGGCGAGGAGGGTCTCTAGTTGCTTCCCTATCTGCAGAGCTTC 480
Db 421 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 481 CTGGAGGTGTGACCGGCTTCTACGCCACCTTGCCACCTTGCCCGAGCCC 522
Db 481 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

## RESULT 7

```
AA004481
ID AA004481 standard; DNA; 525 BP.
XX AC AA004481;
XX 15-NOV-1990 (first entry)
XX G-CSF gene isolated from peripheral blood macrophages.
XX Granulocyte colony stimulating factor; hG-CSF; macrophage;
XX tumour therapy; leukaemia; ss.
XX Homo sapiens.
XX EP272703-A.
XX 29-JUN-1988.
```



Db 148 ctgctcgacactctctgggcatccctgggctccctgagcagctgccccagcagcc 207  
Qy 205 CTGACAGTGGCAGGCTGTGAGCAACTCCATACGCGCCTTTCTCTACCAAGGGGCTC 264  
Db 208 ctgagctggcaggtgtgtgagcaactccatagcgcccttttctctaccaggggtc 267  
Qy 265 CTGACAGGCTTGGAGGATCTCCCGAGTTGGGTCCACCTTGGACACACTGCAGCTG 324  
Db 268 ctgcaaggcccttgaaggatctccccagattggtgtccacccttggacacactgcagctg 327  
Qy 325 GACGTGCGCGACTTTGGCCACCACCATCTGCGCAGCAGATGGAAGAACTGGGAATGCCCTC 384  
Db 328 gacgtcgccgactttgcccaccacactctggcagcagatggaagaactgggaatggccct 387  
Qy 395 GCCCTGACGCCACCCAGGCTGCATGCGCGCTTCGCTTCCAGGCGCGGCA 444  
Db 398 gccctgcaagccaccaggtgccaatgcgcgcttgcctctgtttccagcgccggca 447  
Qy 445 GGAGGGTCTTAGTTGGCTCCCATCTGCAGAGCTTCCCTGGAGGTGCTGACCGCTTCTA 504  
Db 448 ggaagggtcctagtgtcctccatctgcagagcttcttgagggtgtctaccgcttcta 507  
Qy 505 CGCCACCTTGGCCAGCC 522  
Db 508 cgccacttggccagccc 525

## RESULT 9

AAD19771  
ID AAD19771 standard; DNA; 615 BP.  
XX  
AC  
XX  
AC AAD19771;  
DT 18-DEC-2001 (first entry)  
XX  
XX Human hg-CSF gene inserted into plasmid p19CSFm.  
XX  
XX Human; granulocyte colony stimulating factor; hg-CSF; protease;  
KW Factor Xa; kanamycin resistance; endoxylanase signal peptide; ds.  
KW  
XX  
XX Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 88..612  
FT /\*tag= a  
FT /product= "Human hg-CSF protein"  
FT /note= "CDS does not include start codon"  
FT /partial  
XX  
XX WO200173081-A1.  
XX  
XX 04-OCT-2001.  
XX  
XX 31-MAR-2001; 2001WO-KR00549.  
XX  
XX 31-MAR-2000; 2000KR-0017052.  
XX  
XX (ROAD ) KOREA ADV INST SCI & TECHNOLOGY.  
XX  
XX Lee S, Jeong K;  
XX  
XX WPI; 2001-616523/71.  
XX  
XX P-PSDB; AAE12153.  
XX  
XX Recombinant plasmid vector comprising an endoxylanase signal sequence,  
PT human granulocyte colony stimulating factor gene and other components,  
PT when transformed into microorganism useful for preparing the colony  
PT stimulating factor .  
XX  
XX Example 1; Fig 3; 50pp; English.  
XX  
XX The invention relates to an Escherichia coli producing and secreting

CC human granulocyte colony stimulating factor (hg-CSF), more specifically,  
CC to a recombinant plasmid constructed to express secretory hg-CSF in  
CC E. coli, an E. coli transformed with that plasmid to secrete hg-CSF,  
CC and a process for preparing hg-CSF using the transformed hg-CSF. The  
CC recombinant plasmid vector comprises of a kanamycin resistance gene, a  
CC promoter, an endoxylanase signal sequence, a nucleotide sequence coding  
CC for an oligopeptide consisting of 13 amino acids including 6 consecutive  
CC histidine residues and a hg-CSF. E. coli transformed with recombinant  
CC plasmid vector is useful for preparing hg-CSF. The method comprises  
CC culturing the microorganism to obtain a hg-CSF fusion protein and  
CC treating the fusion protein with a protease preferably Factor Xa, to  
CC obtain a hg-CSF, where the fusion protein is obtained from the culture  
CC by employing Ni-column. The present sequence is human hg-CSF gene  
CC inserted into plasmid p19CSFm.  
XX

SQ Sequence 615 BP; 105 A; 219 C; 178 G; 113 T; 0 other;

Query Match 93.3%; Score 486.8; DB 22; Length 615;  
Best Local Similarity 98.6%; Pred. No. 2.4e-101;  
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 25 CTACACACAGAGCTTCCTTTTAAAAAGCTTAGAGCAAGTAGAGAAATCCAGGGCCGATGGC 84  
Db 112 ctgccccagagcttctctcaagtgtcttagagcaagt-gaggaagatccagggcgatg 171  
Qy 85 GCAGCGCTCCAGGAGAGCTGTGCGCACTACAGCTGTGCCACCCGAGAGCTGGTG 144  
Db 172 gcagcgctccaggagaaagtgtgtccacctacaagctgtgccaccgcgagagctggtg 231  
Qy 145 CTGCTCGGACACTCTCTGGGATCCCTGGGCTCCCTCGAGCAGCTGCCAGCCAGGCC 204  
Db 232 ctgctcgagacactctctgggcatccctgggctccctgagcagctgccccagcagcc 291  
Qy 205 CTGACAGTGGCAGGCTGCTTGAGCAACTCCATACCGGCCCTTTCTCTACCAAGGGGCTC 264  
Db 292 ctgcagctggcaggtgtgtgagcaactccatagcggccttttctctaccaggggtc 351  
Qy 265 CTGACAGGCTTGGAGGATCTCCCGAGTTGGGTCCACCTTGCACACACTGCAGCTG 324  
Db 352 ctgcagggccctggaaaggatctcccccgagttggtgtccacctgtgacacactgcagctg 411  
Qy 325 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGAAGAACTGGGAATGCCCTC 384  
Db 412 gacgtcgcgactttgtccaccacactctgagcagatggaagaactgggaatggccct 471  
Qy 395 GCCCTGAGAGCCACCCAGGCTGCATGCGCGCTTCGCTTTCAGGCGCGGGCA 444  
Db 472 gccctgcagccacccaggtgccaatgcgccttcgctctgcttccagcgccggca 531  
Qy 445 GGAGGGTCTTAGTTGGCTCCCATCTGCAGAGCTTCCCTGGAGGTGCTGACCGCTTCTA 504  
Db 532 ggaagggttcctagtgtgctcccatctgcagagcttccctggagggtgctaccgcttcta 591  
Qy 505 CGCCACCTTGGCCAGGCC 522  
Db 592 cgccacttggccagccc 609

## RESULT 10

AAN70223  
ID AAN70223 standard; DNA; 1520 BP.  
XX  
XX  
AC AAN70223;  
XX  
DT 25-APR-1991 (first entry)  
XX  
DE Plasmid pBRV2 insert.  
XX  
KW Human granulocyte colony stimulating factor; G-CSF; leukaemia; ss.  
XX OS Homo sapiens.  
XX

FH Key Location/Qualifiers  
 FT CDS 31..645  
 FT /\*tag= a  
 PN EP220520-A.  
 XX  
 PD 06-MAY-1987.  
 XX  
 PF 30-SEP-1986; 86EP-0113446.  
 XX  
 PR 30-SEP-1985; 85JP-0217150.  
 PR 17-JUL-1986; 86JP-0166710.  
 PR 17-JUL-1986; 86JP-0166709.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Yamazaki T, Nagata S, Tsuchiya M;  
 XX  
 DR WPI; 1987-124182/18.  
 DR P-PSDB; AAP70162.  
 XX  
 PT Polypeptide with human granulocyte colony stimulating factor  
 PT activity - is obtd. by cultivating transformant formed by  
 PT recombinant DNA procedures.  
 XX  
 PS Disclosure; Fig 4; 73pp; English.  
 XX  
 CC The plasmid was isolated from a cDNA library prepd. from CHU-2  
 CC cells, a human oral cavity tumour cell line, using a 1500 bp  
 CC insert from pBRG4 (AAN70222) and probe LC (AAN71351). The plasmid  
 CC was used to prepare recombinant expression plasmids for the prodn.  
 CC of h G-CSF.  
 CC See also AAN70221-N70224 and AAN71349-N71351.  
 XX  
 SQ Sequence 1520 BP; 303 A; 487 C; 402 G; 328 T; 0 other;

Query Match 93.3%; Score 486.8; DB 8; Length 1520;  
 Best Local Similarity 98.6%; Pred. No. 2.7e-101;  
 Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 526  
 QY 25 CTACACAGAGCTTCCCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGCGGATGGC 84  
 DB 145 ctgccccagagcttcctcaagcttagagcaagtgaggagatccagggcgatggc 204  
 QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAGCTGTGCCACCCGAGGAGCTGGTG 144  
 DB 205 gcagcgctccagggagagctgtgtgccacctacaagctgtgccaccccgaggagctgggtg 264  
 QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCGCCAGCGGC 204  
 DB 265 ctgctcgagacactctctggcatccctggctccctgagcagctgccccagcagcc 324  
 QY 205 CTGAGCTGGAGGCTGCTTGGAGCAACTCCATAGAGGCGCTTTCTCTACAGGGGCTC 264  
 DB 325 ctgagctggcaggtgctgtgagcaactccatagagcgctttctctaccaggggctc 384  
 QY 265 CTGAGGCGCTTGGAGGAGTCTCCCGAGTGTGGTCCACCTTGACACACTGCAGCTG 324  
 DB 385 ctgagggccttggagggatctcccccagttggtgtccacacttgacacactgcagctg 444  
 QY 325 GACCTCGCGCACTTTGCCACACCATCTGCAGCAGATGGAAGTGGGAATGGCCCTT 384  
 DB 445 gacgtccgacttggccaccaccttgccagcagatggaagaactgggaatggccct 504  
 QY 385 GCGCTGAGCCACCCAGGAGTGCATGCCGCTTCGCTCTGCTTTCAGGCGCGGCA 444  
 DB 505 gccctgagccacccaggggtgcatgcccgttcgctctgctcttccagcgccggga 564  
 QY 445 GGAGGGTCTAGTGTGCTCCCATCTCAGAGCTTCTGAGAGTGTCTGACGCGTCTTA 504  
 DB 565 ggaaggggtctagttgcttcccatctgacagagcttctggaggtgctgtaccgcttcta 624

QY 505 CGCCACCTTGCCCGAGGCC 522  
 DB 625 cgccaccttgcccgagccc 642  
 RESULT 11  
 AAN71320  
 ID AAN71320 standard; cDNA; 1520 BP.  
 XX  
 AC AAN71320;  
 XX  
 DT 23-APR-1991 (first entry)  
 XX  
 DE Sequence encoding human granulocyte colony stimulating factor (hGCSF)  
 DE in pBRV2.  
 XX  
 KW Leukopenia therapy; neutropenia; eosinopenia; lymphopenia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 31..120  
 FT /\*tag= a  
 FT mat\_peptide 121..645  
 FT /\*tag= b  
 XX  
 PN EP217404-A.  
 XX  
 PD 08-APR-1987.  
 XX  
 PF 03-OCT-1986; 86EP-0113671.  
 XX  
 PR 02-JUN-1986; 86JP-0125660.  
 PR 04-OCT-1985; 85JP-0220450.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Tamura M, Nomura H, Hattori K, Ono M;  
 XX  
 DR WPI; 1987-095431/14.  
 DR P-PSDB; AAP71384.  
 XX  
 PT Leukopenia treating agent esp. for producing fully mature  
 PT neutrophils - contains human granulocyte colony stimulating  
 PT factor obtd. by recombinant DNA methods etc.  
 XX  
 PS Claim 7; Fig 2; 34pp; English.  
 XX  
 CC The hGCSF has the following properties: (i) mol. wgt. 19000 +/- 1000  
 CC (by SDS-PAGE); (ii) isoelectric pt. at 5.5, 5.8 or 6.1 each +/- 0.1;  
 CC (iii) UV max. absorption at 280nm and min. at 250nm. Prodn. of the  
 CC gene, vector etc. is described in JP 269455, 269456, 270838 and  
 CC 270839, each of 1985.  
 XX  
 SQ Sequence 1520 BP; 305 A; 488 C; 401 G; 326 T; 0 other;

Query Match 93.3%; Score 486.8; DB 8; Length 1520;  
 Best Local Similarity 98.6%; Pred. No. 2.7e-101;  
 Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 25 CTACACAGAGCTTCCCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGCGGATGGC 84  
 DB 145 ctgccccagagcttcctcaagcttagagcaagtgaggagatccagggcgatggc 204  
 QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAGCTGTGCCACCCGAGGAGCTGGTG 144  
 DB 205 gcagcgctccagggagagctgtgtgccacctacaagctgtgccaccccgaggagctgtg 264  
 QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCGCCAGCGGC 204  
 DB 265 ctgctcgagacactctctggcatccctggctccctgagcagctgccccagcagcc 324

Qy 205 CTGACGCTGGCAGGCTGCTTGGAGCAACTCCATAGCGCCCTTTCTCTTACACAGGGGCTC 264  
Dd |||||||  
375 ctgcagctggcagcgtctgaagcaactccatagcgcctttctctctacacagggctc 384  
Qy 265 CTGACAGCCCTGGAGGAGTCTCCCGAGTTGGTCCACTTGGACACACTGCAGCTG 324  
Dd |||||||  
385 ctgcagggcctggaaggagctctcccgagtggtggtccacacttggaacacactgcagctg 444  
Qy 325 GACCTGCGCGACTTGGCCACCACTCTGCGAGTGGAGCAAGTGGAGCAAGTGGCCCT 384  
Dd |||||||  
445 gacgtcgccgacttggccaccacatctggcagcagatgggaagactgggaatggccct 504  
Qy 385 GCCCTGACGCGCCACCGAGGTGCCATCGCGGCTTGGCTCTGCTTTCAGCGCGGGCA 444  
Dd |||||||  
505 gccctgcagccacccaggtggcctgcccctgctctgcttccagcgcggca 564  
Qy 445 GGAGGGTCTAGTGGCTCCACTCTGCAGAGCTTCTGGAGGTCTGCTACCGGCTTCTA 504  
Dd |||||||  
565 ggaggggtcctagtgctcccatctgcagagcttctgaggtgctgtaaccggttcta 624  
Qy 505 CGCCACCTTGGCCAGCCC 522  
Dd |||||||  
625 cgcacacttggccagccc 642

## RESULT 12

AA60937  
ID AA60937 standard; DNA; 1521 BP.  
XX  
AC AA60937;  
XX  
DT 18-OCT-1991 (first entry)  
XX  
DE Plasmid pBRV2 insert encoding granulocyte CSF.  
XX  
KW G-CSF; granulocyte colony stimulating factor; antimicrobial.  
XX

## Synthetic.

XX Key Location/Qualifiers  
FH CDS 31..645  
FT mat\_peptide /tag= a  
FT 121  
FT /tag= b  
XX WO8604506-A.  
XX WO8604605-A.  
XX 14-AUG-1986.  
XX 07-FEB-1986; 86WO-JP000053.  
XX 08-FEB-1985; 85JP-0023777.  
XX 03-DEC-1985; 85JP-0270839.  
XX (CHUS ) CHUGAI SEIYAKU KK.  
XX (ONOM/) ONO M.  
XX Ono M, Nomura H, Tamura M, Matsumoto M;  
XX WPI; 1986-225384/34.  
XX P-PSDB; AAP61341.  
XX Infection-protection inducing agent - contains human granulocyte  
PT colony stimulating factor obtd. by cell culture or recombinant  
PT technology.  
XX  
XX Disclosure; Fig 4; 139pp; Japanese.  
XX  
XX The plasmid encodes a novel polypeptide having human granulocyte  
CC colony stimulating factor activity. The CSF may be produced either  
CC from a cell-line secreting the factor into its medium, or from a  
CC transformant microorganism. The product induces high immunity to

CC infection by anaerobic and aerobic bacteria and fungi.  
XX  
SQ Sequence 1521 BP; 304 A; 488 C; 402 G; 327 T; 0 other;

Query Match 93.3%; Score 486.8; DB 7; Length 1521;  
Best Local Similarity 98.6%; Pred. No. 2.7e-101;  
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 25 CTACACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGCG 84  
Dd |||||||  
145 ctgccccagagcttctctcaagtgttagcaagtgaagaaatccagggcgatgac 204  
Qy 85 GCAGCGCTCCAGGAGAGCTGTGCCACCTACAGCTGTGCCACCGAGAGAGCTGTG 144  
Dd |||||||  
205 gcagcgctcccgaggagagctgtgccccaccatcaagctgtgccaccgcgagagctggtg 264  
Qy 145 CTGCTCGACACTCTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCAGCCAGGCC 204  
Dd |||||||  
265 ctgctcgagacactctctgggcatccctgggctccctgagcagctgccccagcagcc 324  
Qy 205 CTGCAGCTGGCAGGCTGCTTGGACCAACTCCATAGCGGCTTTCCTTACACAGGGGCTC 264  
Dd |||||||  
325 ctgcagctggcagcgtctgagccacactccatagcgcctttctctctacacagggctc 384  
Qy 265 CTGCAGGCGCTGGAGGAGTCTCCCGAGTTGGTCCCGCTTGGACACACTGCAGCTG 324  
Dd |||||||  
385 ctgcagccttggagagatctccccaggttgggtccacactggacacactgcagctg 444  
Qy 325 GACGTGCGCGACTTGGCCACCACTCTGGCAGCAGATGGAGCAAGTGGAGTGGCCCT 384  
Dd |||||||  
445 gacgtcgccgacttggccaccacatctggcagcagatgggaagactgggaatggccct 504  
Qy 385 GCCCTGACGCGCCACCGAGGTGCCATCGCGGCTTGGCTCTGCTTTCAGCGCGGGCA 444  
Dd |||||||  
505 gccctgcagccacccaggtggcctgcccctgctctgcttccagcgcggca 564  
Qy 445 GGAGGGTCTAGTGGCTCCACTCTGCAGAGCTTCTGGAGGTCTGCTACCGGCTTCTA 504  
Dd |||||||  
565 ggaggggtcctagtgctcccatctgcagagcttctgaggtgctgtaaccggttcta 624  
Qy 505 CGCCACCTTGGCCAGCCC 522  
Dd |||||||  
625 cgcacacttggccagccc 642

## RESULT 13

AA81478  
ID AA81478 standard; cDNA; 1525 BP.  
XX  
AC AA81478;  
XX  
DT 14-DEC-1990 (first entry)  
XX  
DE Sequence encoding human granulocyte colony-stimulating factor (G-CSF)  
DE in clone pP12 of MIA PaCa cDNA library.  
XX  
KW Leukaemia therapy; anaemia treatment; leukocyte; lymphokine;  
KW interleukin-3; ss.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FH CDS 43..132  
FT /tag= a  
FT mat\_peptide 133..657  
FT /tag= b  
FT conflict 237..238  
FT /tag= c  
FT /note="CHU-2 cDNA clones of Nagata et al. contains a  
FT 9 base pair insertion (GTGATGGAG)".  
FT conflict 588  
FT /tag= d

```
FT conflict /note="G in the CHU-2 clone"
FT 1237 /*tag= e
FT /*tag="C in the CHU-2 clone"
PN EP256843-A.
XX 24-FEB-1988.
XX 11-AUG-1987; 87EP-0307114.
XX 18-NOV-1986; 86US-032037.
PR 01-AUG-1986; 86US-0895194.
XX (CETU ) CETUS CORP.
PI Devlin JJ, Devlin PE, Kawasaki ES, Warren MK;
XX WPI; 1988-051585/08.
DR P-PSDB; AAP81162.
XX Human recombinant granulocyte colony-stimulating factor -
PT causes differentiation of granulocytes and monocytes from bone
PT marrow progenitor cells
XX Example; Fig 4; 61pp; English.
XX The major difference between this clone and the CHU-2 cDNA clones of
CC Nagata et al. (EMBO Journal 5:3,575-581(1986)) are given in FT of
CC AAN81478. The G-CSF was detected in purified mRNA fraction obtd. by
CC sucrose density gradient centrifugation of induced MIA PaCa-2 cells as
CC described in W08804607.
XX Sequence 1525 BP; 303 A; 491 C; 402 G; 329 T; 0 other;
SQ
Query Match 93.3%; Score 486.8; DB 9; Length 1525;
Best Local Similarity 98.6%; Pred. No. 2.7e-101;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 25 CTACACAGAGCTTCCCTTTTAAAGCTTAGAGCAAGTGGAGGAGATCCAGGCGGATGCG 84
DB 157 ctgcccagagcttccctgcacagtgttagagcaagtggagagatccaggcgatggc 216
QY 85 GCAGCGCTCCAGGAGAGCTGTGCCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
DB 217 gcagcgtccaggagaagctgtgcccacatacaagctgtgccacccagagagctggtg 276
QY 145 CTGCTCGGACACTCTCTGGGATCCCTGGGCTCCCTTGAGCAGTGTGCCACCGAGGCC 204
DB 277 ctgctcgagactctctggtccctggctccctgagcagctgccccagcagggcc 336
QY 205 CTGCAGCTGGCAGGCTGTGGAGCACTCCATAGCGGCTTTTCTCTACAGGGGCTC 264
DB 337 ctgcagctggcaggtgctgtgagccactccatagcgcccttctctaccaggggctc 396
QY 265 CTGCAGGCTCGAAGGAGTCTCCCGAGTGTGGGTGCCACCTTGGACACACTGACGCTG 324
DB 397 ctgcagggcttggagggatctcccccagctgggtccacacttgacacactgacgtg 456
QY 325 GACGTGCGGACTTGGCAGACACATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCT 384
DB 457 gacgtcgagacttggcacacacatctggcagcagatggaagaactgggaatggccct 516
QY 385 GCCTTGAGCCACCCAGGAGTGCATGCCGCTTGCCTTCTCTACAGCGCGGGA 444
DB 517 gacctgagccccacaggggtgcaatgcccgtctgctctgcttccagcgccggga 576
QY 445 GGAGGGTCTTAGTGGCTCCCATCTGCAGAGCTTCTGGAGGTGTGCTACCGCTCTTA 504
DB 577 ggaaggggtcctagtgtcctcccatctgcagagcttctctgaggtgctgtaaccggtcta 636
QY 505 CGCCACCTTGCCCGAGCCC 522
```

```
DB 637 cgccaccttgcccagccc 654
RESULT 14
AAN91086
ID AAN91086 standard; DNA; 1525 BP.
XX AAN91086;
AC AAN91086;
XX 04-JUL-1990 (first entry)
XX Plasmid pP12 contg. colony stimulating factor-1 gene.
XX Colony stimulating factor-1; pP12; ss.
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
FH sig_peptide 43..132
FT /*tag= a
FT mat_peptide 133..654
FT /*tag= b
XX W08901038-A.
XX PN
XX 09-FEB-1989.
XX PD
XX 20-JUL-1988; 88WO-US02445.
XX PF
XX 24-JUL-1987; 87US-0077188.
XX PR
XX (CETU ) CETUS CORP.
XX PA
XX Kawasaki ES, Devlin JJ, Martin G, O'Rourke E, Clark R;
XX WPI; 1989-061174/08.
XX DR P-PSDB; AAP95033.
XX PT Recombinant Bacovirus transfer vectors - used for prodn. of colony
XX stimulating factor in Baculovirus insect cell expression system.
XX PS Disclosure; Fig 3; 68pp; English.
XX The cDNA insert of pP12 contains 11 more bases than the CHU-2 G-CSF
CC clone. The major difference between this clone, derived from MIA
CC PaCa-2, and CHU-2 clone is a 9 bp insertion in the CHU-2 clone between
CC bases 237 and 238. There are two other differences: an A at position 588
CC (G in the CHU-2 clone) is a silent third base change, and a T at
CC position 1237(C in the CHU-2 clone) in the 3' untranslated region.
CC See also AAN91085.
XX SQ Sequence 1525 BP; 303 A; 490 C; 403 G; 328 T; 1 other;
Query Match 93.3%; Score 486.8; DB 10; Length 1525;
Best Local Similarity 98.6%; Pred. No. 2.7e-101;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 25 CTACACACAGCTTCCCTTTTAAAGCTTAGAGCAAGTGGAGGATCCAGGCGGATGCG 84
DB 157 ctgcccagagcttccctgcacagtgttagagcaagtggagagatccaggcgatggc 216
QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTG 144
DB 217 gcagcgtccaggagaagctgtgcccacatacaagctgtgccacccagagagctggtg 276
QY 145 CTGCTCGGACACTCTCTGGGATCCCTGGGCTCCCTTGAGCAGTGTGCCACCGAGGCC 204
DB 277 ctgctcgagactctctggtccctggctccctgagcagctgccccagcagggcc 336
QY 205 CTGCAGCTGGCAGGCTGTGGAGCACTCCATAGCGGCTTTTCTCTACAGGGGCTC 264
```

Db 337 ctgcagctggtgagcagctgttgagcgaactccatagcggtcttcttcttaccaggggtc 396  
Qy 265 CTGAGGCGCTTGAAGGATCTCCCGAGTTGGTCCACCTTGACACACTGCAGCTG 324  
Db 397 ctgcagccttgaaggtatctcccgagttggttccaccttgacacactgcagctg 456  
Qy 325 GACGTGCGCGACTTTGCCACCACTATCTGCGAGCAGATGGAAGAACTGGGAATGCCCT 384  
Db 457 gacgtcgcgactttgcccaccacctctggcagcagatggagaactgggaatggccct 516  
Qy 385 GCCCTGAGCGCCACCGAGGNGCCATGCCGCTTCGCTCTGCTTCCAGCGCGGCA 444  
Db 517 gccctgcagccaccacccgggtgcctatgcggccttcgctctgtttccagcgccgga 576  
Qy 445 GGAGGGTCTCTAGTTGGCTCCCATCTGCAGAGCTTCTGAGGTTCTGATCCGCTTCTA 504  
Db 577 ggaaggtctctagttgctcccatctgcagagcttctgcaggtgtctgtacccggttcta 636  
Qy 505 CGCCACCTTGCCGAGGCC 522  
Db 637 cgcacacttgccagccc 654

## RESULT 15

AAN71089  
ID AAN71089 standard; DNA; 1415 BP.  
XX AC AAN71089;  
XX AC AAN71089;  
DT 26-APR-1991 (first entry)  
XX DE Sequence of human granulocyte colony stimulating factor (hpG-CSF)  
DE cDNA clone Ppo2.

XX Haematopoietic disorders; therapy; aplastic anaemia;  
KW bone marrow transplant; burn wounds; leukaemia; ds.  
XX Homo sapiens.

XX Key Location/Qualifiers  
FH CDS 1..36  
FT /\*tag= a  
FT mat\_peptide 37..561  
FT /\*tag= b  
FT polyA\_signal 1797..1802  
FT /\*tag= c

PN WO8701132-A.

XX PD 26-FEB-1987.

XX PF 22-AUG-1986; 86WO-U001708.

XX PR 03-MAR-1986; 86US-0835548.

XX PR 23-AUG-1985; 85US-0768959.

XX PR 23-AUG-1985; 85US-0768954.

XX (KIRI-) KIRIN-AMGEN INC.

PA (KIRI ) KIRIN-AMGEN INC.

XX Souza LM;

XX WPI; 1987-064855/09.

XX P-PSDB; AAP70730.

XX Poly:peptide with granulocyte colony stimulating factor activity  
PT - obtd. by recombinant DNA procedures for treating haematopoietic  
PT disorders

XX Disclosure; Page 22-24; 79pp; English.

XX The examples describe procedures for the designing of probes for  
CC hpG-CSF cDNA and genomic clones, both of which are claimed.

CC Specifically claimed are DNA sequences encoding for (Ala 1)hpG-CSF;  
CC (Ser 36, 42, 64 and 74)hpG-CSF and the corresponding Met-1 cpds.  
CC The novelty is that hpG-CSF is the prod. of procarvotic or  
CC eucaryotic expression of an exogenous DNA sequence. The construction  
CC of hpG-CSF expression vectors is also described in the examples.

XX SQ Sequence 1415 BP; 265 A; 451 C; 380 G; 319 T; 0 other;

Query Match 93.0%; Score 485.2; DB 8; Length 1415;  
Best Local Similarity 98.4%; Pred. No. 6.2e-101;  
Matches 490; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 25 CTACACAGAGCTTCTCTTTTAAAAAGCTTAGAGCAAGTGAGGAAGATCCAGGGCGATGCG 84  
Db 61 ctgccccagagcttctgctcaagtgtcttagagcaagtggaggaagatccagggcagatgac 120  
Qy 85 CGAGGCTCCAGGGAAGCTGTGTGCCACTACAGCTGTGCCACCGCCGAGAGCTGGTG 144  
Db 121 gcagcgtccaggaagaagtgtgtgccacctacaagctgtgccaccccgagagctgtg 180  
Qy 145 CTGCTCGGACACTCTCTTGGCATCCCTGGGCTCCCTGAGCAGCTGCCACAGCCAGGCC 204  
Db 181 ctgctcggacactctctctgggcatccccctgggtccctctgagcagctgccccagcagccc 240  
Qy 205 CTGACAGCTGGCAGGCTGCTTGAGCCAACTCCATAGCGGCTTTTCTCTTACAGGGGCTC 264  
Db 241 ctgcagctggcagctgctgtgagccaaactccatagcgccctttctctaccaggggtc 300  
Qy 265 CTGACAGGCGCTTGAAGGATCTCCCGAGTTGGTGGTCCACCTTGGACACACTGCAGCTG 324  
Db 301 ctgcagggcctggaaggatctcccccgagttgggtccacacctggacacactgcagctg 360  
Qy 325 GACGTGCGCGACTTTTGGCCACCACTCTGGCAGCAGATGGGAAGAACTGGGAATGCCCT 384  
Db 361 gacgtcgcgactttgccccaccacctctggcagcagatgggaagaactgggaatggccct 420  
Qy 385 GCCCTGACAGCCACCCAGGGTGCCATGCCGCTTTCGCTCTGTGTTTCCAGCGCGGCA 444  
Db 421 gccctgcagccaccacccaggtgcatgccccttcgctctgcttccagcgccggca 480  
Qy 445 GGAGGGTCTCTAGTTGGCTCCCATCTGCAGAGCTTCTGAGGTTCTGATCCGCTTCTA 504  
Db 481 ggaggggtcctggtgtgctcccatctgcagagcttctctggaggtgtcgtacccggttcta 540  
Qy 505 CGCCACCTTGCCGAGGCC 522  
Db 541 cgcacacttgccagccc 558

Search completed: May 8, 2002, 12:28:35  
Job time: 12566 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2002, 07:52:27 ; Search time 1794.34 Seconds  
(without alignments)

3926.464 Million cell updates/sec

Title: US-09-680-514-4\_COPY\_526\_1047

Perfect score: 522

Sequence: 1 GGCCCAACATATCGCGCTC.....TACGCCACCTTGCCCGACCC 522

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*  
2: em\_esthum:\*\*  
3: em\_estin:\*\*  
4: em\_estmu:\*\*  
5: em\_estov:\*\*  
6: em\_estpl:\*\*  
7: em\_estro:\*\*  
8: em\_hic:\*\*  
9: gb\_est1:\*\*  
10: gb\_est2:\*\*  
11: gb\_hic:\*\*  
12: gb\_gss:\*\*  
13: em\_gss\_hum:\*\*  
14: em\_gss\_inv:\*\*  
15: em\_gss\_pln:\*\*  
16: em\_gss\_vrt:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	454.2	87.0	990	10	BM423896
2	387.8	74.3	817	10	BM009358
3	343.6	65.8	570	10	BI961242
4	323.6	62.0	548	10	BI960812
5	320.4	61.4	973	10	BI411128
6	312.8	59.9	598	10	BG548320
7	300.6	57.6	507	10	BE485194
8	281.2	53.9	829	10	BM009247
9	210.6	40.3	948	10	BI822673
10	181	34.7	368	10	BI961002
11	150	28.7	509	10	BM256261
12	135.8	26.0	327	10	BF848766
13	67.6	13.0	673	10	BI961882
14	54.8	10.5	925	12	CNS0091P
15	50.2	9.6	513	10	BE480590
16	49	9.4	925	12	CNS0091P
17	48.8	9.3	617	9	AI981598

c 18	43.2	8.3	461	10	BE598945
c 19	43.2	8.3	559	10	BG556176
c 20	42.6	8.2	452	9	AA155632
c 21	42.6	8.2	899	12	CNS01VYT
c 22	42.4	8.1	1203	12	CNS015Y4
c 23	42.2	8.1	598	9	AA115932
c 24	41.4	7.9	421	9	AA058743
c 25	41.4	7.9	478	10	BI345642
c 26	41.2	7.9	676	9	BB647049
c 27	40.8	7.8	545	10	BF868325
c 28	40.8	7.8	893	10	BM016196
c 29	40.8	7.8	1011	10	BM474138
c 30	40.6	7.8	447	10	BG604563
c 31	40.6	7.8	845	9	AL572931
c 32	40.6	7.8	932	12	CNS0072Q
c 33	40.6	7.8	1003	10	BI410408
c 34	40.4	7.7	458	9	AW273202
c 35	40.4	7.7	485	10	BE463718
c 36	40.2	7.7	864	12	CNS022LE
c 37	40	7.7	368	10	BI961002
c 38	40	7.7	548	10	BI960812
c 39	40	7.7	570	10	BI961242
c 40	39.8	7.6	387	9	AW193322
c 41	39.8	7.6	427	9	BE048584
c 42	39.8	7.6	432	9	AW175624
c 43	39.8	7.6	441	9	AI818468
c 44	39.8	7.6	444	9	AW072844
c 45	39.8	7.6	450	9	AW273147

#### ALIGNMENTS

RESULT 1  
BM423896  
LOCUS AGENCOURT\_6399123 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5517129  
DEFINITION 5', mRNA sequence.  
ACCESSION BM423896  
VERSION BM423896.1 GI:18392108  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 990)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LICM2018 row: n column: 10  
High quality sequence stop: 542.  
Location/Qualifiers  
1. 990

#### FEATURES

Source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5517129"  
/clone\_lib="NIH\_MGC\_41"  
/tissue\_type="amelanotic melanoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pOPB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California,

BE598945 P11\_84\_E1  
BG556176 EMI\_68\_F0  
AA155632 z070d02\_s  
AL169742 Tetraodon  
AL106054 Drosophil  
AA115932 z101a10\_s  
AA058743 zk70f07.s  
BI345642 374523 MA  
BB647049 BB647049  
BF868325 292084 BA  
BM016196 603642649  
BM474138 AGENCOURT  
BG604563 WHE0947\_C  
AL572931 AL572931  
AL066742 Drosophil  
BI410408 602964928  
BI960812 MONOL\_1\_B  
BI961242 MONOL\_7\_G  
AW193322 xl73e08.x  
BE048584 hr50h10.x  
AW175624 RC3-BT004  
AI818468 wk60c03.x  
AW072844 xa42q06.x  
AW273147 xr34c05.x



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

REFERENCE 1 (bases 1 to 570)  
AUTHORS Vandenplas, M.L., Cordonnier-Pratt, M.-M., Sudman, M.L., Wentzel, V.E.,  
Gingle, A.R., Pratt, L.H. and Moore, J.N.  
TITLE An EST database from equine (Equus caballus) monocytes  
JOURNAL Unpublished (2001)  
COMMENT Contact: Cordonnier-Pratt MM  
Department of Botany  
The University of Georgia  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 542 1805  
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude polyA, vector and regions  
below Phred quality 16. The threshold for high quality sequence is  
20. Three-prime sequences, which are obtained with PolyMix or T7  
sequencing primer, are presented as the reverse complement.

Seq primer: JEN REV

High quality sequence stop: 541

POLYA-No.

FEATURES  
Source

Location/Qualifiers  
1..570

/organism="Equus caballus"

/db\_xref="taxon:9796"

/clone\_lib="Monocytes (MONOL)"

/cell\_type="Isolated peripheral blood monocytes stimulated  
with E. coli lipopolysaccharide"

/note="Vector: pBluescript SK(-) from Lambda ZapII;

Site\_1: XhoI; Site\_2: EcoRI; The library was made from

poly-A RNA in the cloning vector lambda ZAPII. Clones to

be sequenced were prepared by mass excision."

BASE COUNT 100 a 202 c 167 g 101 t

ORIGIN

Query Match 65.8%; Score 343.6; DB 10; Length 570;

Best Local Similarity 87.4%; Pred. No. 8.5e-68;

Matches 376; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 25 CTACCACAGAGCTTCCTTTTAAAGAGCTTAGAGCAAGTGAGGAAGATCCAGGCCGATGGC 84

DB 141 CTGCCCCAGAGCTTCCTCTCAAGTGCCTTAGAGCAAGTGAGGAAGATCCAGGCCGATGGT 200

QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCGAGGAGCTGGTG 144

DB 201 GCAGCGCTGCAGGACAGCTGTGTGCCACCCACAAAGCTGTGCCACCCCTCAGGAGCTCATG 260

QY 145 CTGCTCGGACACTCTCTCTGGGCATCCCTGTGGCTCCCTTGAGCAGCTGCCCCAGCCAGGCC 204

DB 261 CTGCTGGGACACTCTCTCTGGGCATCCCTGAGCAGCTCTCTTCCCTGAGCAGCTGTCTCCAGCCAGGCC 320

QY 205 CTGAGCTGGCAGGCTGCTTGAGCCAACTTCAATAGCGGCCCTTTTCTCTACCAAGGGGCTC 264

DB 321 CTGAGCTGGCAGGCTGCTTGAGCCAACTTCAATAGCGGCCCTTTTCTCTACCAAGGGGCTC 380

QY 265 CTGAGGCGCTTGGAGGAGTCTCCCGAGTTGGGTGGGTGCCACCTTGACACACTGCAGCTG 324

DB 381 CTGAGGCGCTTGGAGGAGTCTCCCGAGTTGGGTGGGTGCCACCTTGACACACTGCAGCTG 440

QY 325 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGGAAGACTGGGAATGCCCCCT 384

DB 441 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGGAAGACTGGGAATGCCCCCT 500

QY 385 GCCCTGAGGCCACCGAGGTGCCATGCGGCTTTCGGCTCTGCTTTCAGGCGCGGCA 444

DB 501 GTGTGACAGCCACCGATGGGCCCATGTCGACCTTCGCTCGGCTTCCAGCGCGGCA 560

QY 445 GGAGGGTCC 454

DB 561 GGAGGGTGC 570.

RESULT 4  
BI960812

LOCUS MONOL\_1\_B08.bl\_A005 Monocytes (MONOL) Equus caballus CDNA, mRNA

DEFINITION sequence.

ACCESSION BI960812

VERSION BI960812.1

KEYWORDS GI:16319015

SOURCE horse.

ORGANISM Equus caballus

REFERENCE 1 (bases 1 to 548)

AUTHORS Vandenplas, M.L., Cordonnier-Pratt, M.-M., Sudman, M.L., Wentzel, V.E.,  
Gingle, A.R., Pratt, L.H. and Moore, J.N.

TITLE An EST database from equine (Equus caballus) monocytes

JOURNAL Unpublished (2001)

COMMENT Contact: Cordonnier-Pratt MM  
Department of Botany  
The University of Georgia  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 542 1805  
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude polyA, vector and regions  
below Phred quality 16. The threshold for high quality sequence is  
20. Three-prime sequences, which are obtained with PolyMix or T7  
sequencing primer, are presented as the reverse complement.

Seq primer: JEN REV

High quality sequence stop: 481

POLYA-No.

FEATURES  
Source

Location/Qualifiers  
1..548

/organism="Equus caballus"

/db\_xref="taxon:9796"

/clone\_lib="Monocytes (MONOL)"

/cell\_type="Isolated peripheral blood monocytes stimulated  
with E. coli lipopolysaccharide"

/note="Vector: pBluescript SK(-) from Lambda ZapII;

Site\_1: XhoI; Site\_2: EcoRI; The library was made from

poly-A RNA in the cloning vector lambda ZAPII. Clones to

be sequenced were prepared by mass excision."

BASE COUNT 98 a 203 c 147 g 98 t

ORIGIN

Query Match 62.0%; Score 323.6; DB 10; Length 548;

Best Local Similarity 86.4%; Pred. No. 2.8e-63;

Matches 356; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 25 CTACCACAGAGCTTCCTTTTAAAGAGCTTAGAGCAAGTGAGGAAGATCCAGGCCGATGGC 84

DB 137 CTGCCCCAGAGCTTCCTGCTCAAGTGCCTTAGAGCAAGTGAGGAAGATCCAGGCCGATGGT 196

QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAAGCTGTGCCACCCGAGGAGCTGGTG 144

DB 197 GCAGCGCTGCAGGACAGCTGTGTGCCACCCACAAAGCTGTGCCACCCCTCAGGAGCTCATG 256

QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGTGGCTCCCTTGAGCAGCTGCCCCAGCCAGGCC 204

DB 257 CTGCTGGGACACTCTCTGGGCATCCCTGAGCAGCTGTCTTCCCTGAGCAGCTGTCTCCAGCCAGGCC 316

QY 205 CTGAGCTGGCAGGCTGCTTGAGCCAACTTCCATACGCGCCTTTTCTCTTACCAAGGGGCTC 264

DB 317 CTGAGCTGGCAGGCTGCTTGAGCCAACTTCCATACGCGCCTTTTCTCTTACCAAGGGGCTC 376

QY 265 CTGAGGCGCTTGGAGGAGTCTCCCGAGTTGGGTGGGTGCCACCTTGACACACTGCAGCTG 324

DB 377 CTGAGGCGCTTGGAGGAGTCTCCCGAGTTGGGTGGGTGCCACCTTGACACACTGCAGCTG 436

QY 325 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGGAAGACTGGGAATGCCCCCT 384

DB 437 GACGTGCGCGACTTTGGCCACCACCATCTGGCAGCAGATGGGAAGACTGGGAATGCCCCCT 496

Qy	197	GCCAGGCCCTGCAGCTGGCAGGCTGTGTAGACCAACTCCATAGCGGCTTTTCTCTACC	256
Dd	271	GCCAGGCCCTGCAGCACACACAGTGCCCTAAAGCCAGCTCCACAGTGGGCTCTGCCTCTACC	330
Qy	257	AGGGCTCTCTGCAGGCCCTGGAAGGGATCTCCGCCGAGTTGGGTCCCACCTTGGACACAC	316
Dd	331	AAGGTCTCTGCAGGCTCTATCGGGTAATTTCCTCTGCCCTGGCCCCCACCTTGGACTTGC	390
Qy	317	TGCAGCTGGAGCTGCGCCGACTTTGCCACCACCCTCTCTGGCAGCAGATGGAAGAAGCTGGGAA	376
Dd	391	TTCACTGGATGTGTCACAATTTGCCACCACCCTCTGGCAGCAGATGGAAGAAGCTAGGGG	450
Qy	377	TGGCCCTCTGCCCTGCAGCCACCCAGGAGGTGCATGCGCGGCTTCGCTCTGCTTTCACGC	436
Dd	451	TGGCCCTCTGCTGCAGCCACACAGACGCGCATGCCAGGCTTCACTTCTTCGCTTCCAGC	510
Qy	437	GCCGGCAGAGGGGCTCTAGTTGCCCTCCACTGTCAGAGCTTCCCTGAGGTGCTGATCC	496
Dd	511	GCCGGCAGAGGGTGTCTCTGGCCATTTCTGTAACCTTGCAGGCTTCTCTGGAGACGCTCGCC	570
Qy	497	CGCTTCTACGCCACTTGGCCAGGCC	522
Dd	571	TTGCTCTGCACCACTTGGCCCTAGACC	596
RESULT 6			
BG548320			
LOCUS 602575289F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703159 5', mRNA linear EST 04-APR-2001			
DEFINITION mRNAs sequence.			
ACCESSION BG548320			
VERSION 1 (bases 1 to 598)			
KEYWORDS NIH-MGC http://mgc.nci.nih.gov/			
SOURCE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)			
ORGANISM Human.			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE 1 (bases 1 to 598)			
AUTHORS NIH-MGC http://mgc.nci.nih.gov/			
TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)			
JOURNAL Contact: Robert Strausberg, Ph.D.			
COMMENT Email: cgab@r@mail.nih.gov			
Tissue Procurement: CLONTECH Laboratories, Inc.			
cDNA Library Preparation: CLONTECH Laboratories, Inc.			
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLCMI541 row: b column: 24			
High quality sequence stop: 597.			
FEATURES			
Location/Qualifiers			
1..598			
Source			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:4703159"			
/clone_lib="NIH_MGC_77"			
/lab_host="DH10B (T1 phage-resistant)"			
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site:1: SfiI (ggcgctcgcc); Site_2: SfiI (ggccattagcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCAATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."			
BASE COUNT 113 a 202 c 163 g 120 t			
ORIGIN			

```
Query Match          59.9%; Score 312.8; DB 10; Length 598;
Best Local Similarity 99.4%; Pred. No. 7.9e-61;
Matches 314; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 207 GCAGCTGCAGGCTGCTTTGAGCAACTCATAGCGGCTTTTCCCTCTACCGGGGCTCTCT 266
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 185 GAAGCTGGCAGGCTGCTTTGAGCAACTCATAGCGGCTTTTCCCTCTACCGGGGCTCTCT 244
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 267 GCAGGCTTGAAGGATCTCCCGAGTGGGTCCCGACCTTGGACACTGCACCTGGA 326
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 245 GCAGGCTTGAAGGATCTCCCGAGTGGGTCCCGACCTTGGACACTGCACCTGGA 304
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 327 CGTCCCGACTTTGCCACCACTCGCAGCAGATGGAAGAACTGGGAATGGCCCTTCG 386
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 305 CGTCCCGACTTTGCCACCACTCGCAGCAGATGGAAGAACTGGGAATGGCCCTTCG 364
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 387 CCTGCAGCCACCCAGGTCATGCCGGCTTCGCCCTCTCTCTTCACGCGCGGCGAGG 446
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 365 CCTGCAGCCACCCAGGTCATGCCGGCTTCGCCCTCTCTCTTCACGCGCGGCGAGG 424
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 447 AGGGTCTAGTTCCTCCCATCGCAGCTTCCTGGAGGTGCTGACCGGTTCTACG 506
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 425 AGGGTCTAGTTCCTCCCATCGCAGCTTCCTGGAGGTGCTGACCGGTTCTACG 484
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 507 CCACCTTCCCGAGCCC 522
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 485 CCACCTTCCCGAGCCC 500
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
LOCUS BE485194 507 bp mRNA linear EST 28-AUG-2000
DEFINITION 172130 BARC 5BOV Bos taurus cdna 5', mRNA sequence.
ACCESSION BE485194
VERSION BE485194.1 GI:9604727
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 507)
AUTHORS Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and
Wells,K.D.
TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary
gland cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tadselpsi@ars.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTAGTCACGACG
Plate: 134 row: J column: 17
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1. 507
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 5BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."
```

```
BASE COUNT 97 a 170 c 158 g 82 t
ORIGIN

Query Match          57.6%; Score 300.6; DB 10; Length 507;
Best Local Similarity 84.1%; Pred. No. 4.3e-58;
Matches 339; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 105 GGTGCCACCTTACAGCTGTGCCACCCGAGGAGCTGGTGTCTCGGACACTCTCTGGG 164
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 104 GGTGCCGCCCCAAGCTGTGCCACCCGAGGAGCTGTGCTCTCAGCACTCTCTGGG 163
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 165 CATCCCTCCGGCTCCCTCGAGCAGCTGCCAGCGCCCTGCAGCTGCAGGCTGTCT 224
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 164 CATCCCTCCAGGCTCCCTTAAGCAGCTGTCTCCAGCAGTCCCTGCAGCTGCCT 223
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 225 GAGCCAACCTCCATAGAGGGCTTTTCTCTACAGGGGCTCTCGAGGCCCTGGAAGGAT 284
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 224 GAACCAACCTACAGCGGGCTCTTTCTCTACAGGGCTCTCTCGAGGCCCTGGCGGCAT 283
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 285 CTCCTCCGAGTTGGTCCACCTTGGACACACTGCAGCTGGAGCTCGCGCACTTTGCCAC 344
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 284 CTCCTCAGAGCTGGCCCTTGGACACACTGCAGCTGGAGCTACCTAGCTTTGCCAC 343
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 345 CACCATCTGGCAGCAGATGGGAATGGCCCTGCCCTGCAGCCACCCAGG 404
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 344 GAACATCTGGCTGCAGATGGAGGACCTGGGGCGGCCCTGTGCAGCCACCCAGG 403
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 405 TGCATGCGCGGCTTCGCTCTGCTTCAGCGCGGGGAGGAGGCTTCAGTTGCCTC 464
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 404 CGCCATGCGGACCTTCACTTCAGCCTTCACGACGAGAGAGGAGGCTCTGGTGTCTTC 463
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 465 CCATCTGCAGACTCTCTGGAGGTGCTGCTACCGCTTCTACGC 507
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 464 CCAGCTGCATCTGTTCTCTGGAGCTGGCATACCGTGGCCTGCSC 506
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
LOCUS BM009247 829 bp mRNA linear EST 30-OCT-2001
DEFINITION 603629663F1 NIH_MGC_41 Homo sapiens cdna clone IMAGE:5434958 5',
mRNA sequence.
ACCESSION BM009247
VERSION BM009247.1 GI:16523601
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 829)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1912 row: n column: 15
High quality sequence start: 7
High quality sequence stop: 590.
Location/Qualifiers
1. 829
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5434958"
/tissue_lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
```



[illegible]

```

DEFINITION OVO-EN0102-081100-458-g06 EN0102 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF848766
VERSION BF848766.1 GI:12235903
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
REFERENCE 1 (bases 1 to 327)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1-QV0&t2-QV0-EN0102-
081100-458-g06&t3-2000-11-08&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 327.
FEATURES
source .
1..327
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0102"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 61 a 87 c 103 g 76 t
ORIGIN
Query Match 26.0%; Score 135.8; DB 10; Length 327;
Best Local Similarity 89.6%; Pred. No. 7e-21;
Matches 146; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 201 GGCCCTGCAGCTGGCAGCTGTTGAGCCCACTCCATAGCGGCTTTCTCTACACAGG 260
Db 139 GGCCCATACACAGGCGCGCTGTTTGAGCAACTCCATAGCGGCTTTCTCTACACAGG 198
QY 261 GCTCTGAGCCCTGGAGGATCTCCCGAGTTGGTCCCACTTGGACACTGCA 320
Db 199 GCTCTGAGCCCTGGAGGATCTCCCGAGTTGGTCCCACTTGGACACTGCA 258
QY 321 GCTGAGCTGCCGACTTTGGCCACCACTCTGGCAGCAGATG 363
Db 259 GCTGAGCTGCCGACTTTGGCCACCACTCTGGCAGCAGTG 301

RESULT 13
LOCUS BI961882
DEFINITION MON01_7_g12_g1_A005 Monocytes (MON01) Equus caballus cDNA, mRNA
sequence.
ACCESSION BI961882

```

```

VERSION BI961882.1 GI:16320085
KEYWORDS EST.
SOURCE horse.
ORGANISM Equus caballus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE 1 (bases 1 to 673)
AUTHORS Vandeplass,M.L., Cordonnier-Pratt,M.-M., Sudman,M.L., Wentzel,V.E.,
Gingle,A.R., Pratt,L.H. and Moore,J.N.
TITLE An EST database from equine (Equus caballus) monocytes
JOURNAL Unpublished (2001)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below phred quality 16. The threshold for high quality sequence is
20. Three-prime sequences, which are obtained with Polymix or T7
sequencing primer, are presented as the reverse complement.
Seq primer: Polymix
High quality sequence start: 39
High quality sequence stop: 673
POLYA=No.
FEATURES
source .
1..673
Location/Qualifiers
/organism="Equus caballus"
/db_xref="taxon:9796"
/clone_lib="Monocytes (MON01)"
/cell_type="Isolated peripheral
with E. coli lipopolysaccharide"
/note="Vector: pBluescript SK(-) from Lambda ZapII;
Site_1: XhoI; Site_2: EcoRI; The library was made from
poly-A RNA in the cloning vector lambda ZapII. Clones to
be sequenced were prepared by mass excision."
BASE COUNT 147 a 178 c 166 t
ORIGIN
Query Match 13.0%; Score 67.6; DB 10; Length 673;
Best Local Similarity 83.5%; Pred. No. 2.2e-05;
Matches 76; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 432 CCAGCGCCGGCAGGAGGGTCTAGTTGCTCCACTGCGAGCTTCCTGGAGGTCTC 491
Db 1 CCAGCGCCGGCAGGAGGGTCTAGTTGCTCCACTGCGAGCTTCCTGGAGGTCTC 60

QY 492 GTACCGCGTTCTACGCCACCTTGCCAGCCC 522
Db 61 ATACCGTGGTCTGCGCTACCTGCGCGAGCCC 91

RESULT 14
LOCUS CNS00091P
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPI-98 library from Drosophila melanogaster (fruit
fly); genomic survey sequence.
ACCESSION AL053013
VERSION AL053013.1 GI:4934461
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

```

COMMENT - Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamooser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

# FEATURES

source  
1..925  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone\_lib="RPCI-98"  
/clone="BACR19D16"  
/note="end : tET3"

BASE COUNT 120 a 61 c 61 g 172 t 511 others  
ORIGIN

Query Match 10.5%; Score 54.8; DB 12; Length 925;  
Best Local Similarity 15.5%; Pred. No. 0.019;  
Matches 66; Conservative 174; Mismatches 182; Indels 3; Gaps 1;  
QY 30 ACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGGCGCAGC 89  
DB 432 ANANNNTTATTATTAAANNANNANANANANNANNNAGCSMCGKCGSTTBGSTTT 551  
QY 90 GCTCCAGGAGAGCTGTGTGCCACCTACAAGCTGTGCCACCCGAGGAGCTGGTGCTCT 149  
DB 552 TTTSSGSGYGGKCSGSGBSGCCSCSSCSGCCSCCSCCSCSSYCCSSBSBSKCS 611  
QY 150 CGGACACTCTCGGCATCCCTCGGCTCCCTGAGCAGCTCCCCAGCCAGCCGCTGCA 209  
DB 612 STBSGCCSCSKSVCGTSSSSSSCSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 671  
QY 210 GCTGCAGGCTCTTGAGCAACTCCATAGCGGCTTTCTCTCTACCAAGGGCTCCTGCA 269  
DB 672 SKSTSASGSGWSAGGSGSTGTSTSSSSSSTSTSSSSVSGSKSTBSGSGBSGSGS 731  
QY 270 GCCTCGGAAGGATCTCCCGAGTTGGTCCCACTTGGACACACTGCA---GCTGGA 326  
DB 732 SSSSTSSBBSCTSTSSSSSSSYSSCTCCCTCCSYSSSTSSSTSSSTSSSTSSSTSS 791  
QY 327 GTCGCCGACTTGGCACCATCTGGCAGCAGATGGAAGACTGGGAATGCCCTCC 386  
DB 792 VETSSSDSTSTCCSCCYMTCTCCCTBMBCTSTSCGSSSSSGKGGVTKCGCGCGSS 851  
QY 387 CTTGAGCCACCCAGGGTGGCATCCGCGCTTCGCTCTTCCAGCGCCGCGCAGG 446  
DB 852 STNGMBGTSSACSSSSSSSSSVSSSSKSSASSSVSSGSGVSSNSSSASKSSSGS 911  
QY 447 AGGGG 451  
DB 912 VSSGS 916

RESULT 15  
BE480590 LOCUS BE480590 513 bp mRNA linear EST 28-AUG-2000  
DEFINITION 165716 BARC 5BOV Bos taurus CDNA 5', mRNA sequence.  
ACCESSION BE480590  
VERSION BE480590.1 GI:9600123  
KEYWORDS EST.  
SOURCE cow.  
ORGANISM Bos taurus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
AUTHORS 1 (bases 1 to 513)  
TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary  
Wellis,K.D.  
JOURNAL Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and  
gland cDNA library  
COMMENT Unpublished (2000)  
Contact: Sonstegard TS  
USDA, ARS, Beltsville Agricultural Research Center  
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA  
Tel: 301 504 8416  
Fax: 301 504 8414  
Email: tads@psi.barc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred 18  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.

PCR Primers  
FORWARD: AGGAAACAGCTATGACCAT  
BACKWARD: GTTTCAGCTCAGCAGC  
Plate: 17 row: C column: 20  
Seq primer: ATTAGGTGACACTATAG.

# FEATURES

Location/Qualifiers  
1..513  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="BARC 5BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"

/note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;  
Library made from pooled mRNA isolated from mammary  
tissues at eight physiological, developmental, and disease  
states."  
BASE COUNT 126 a 150 c 130 g 107 t  
ORIGIN

Query Match 9.6%; Score 50.2; DB 10; Length 513;  
Best Local Similarity 81.7%; Pred. No. 0.18; Indels 0; Gaps 0;  
Matches 58; Conservative 0; Mismatches 13;  
QY 25 CTACCACAGAGCTTCCTTTTAAAGCTTAGAGCAAGTGAAGAGATCCAGGGCGATGC 84  
DB 443 CTGCCCCAGAGCTTCTGCTCAAGTGCTTAGAGCAAGTGAAGAAATCCAGGCTGATGC 502  
QY 85 GCAGCGCTCCA 95  
DB 503 GCCGAGCTGCA 513

Search completed: May 8, 2002, 09:29:56  
Job time: 5849 sec

---



PI ITOU SEIGA  
PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC  
C12N15/27,  
PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,  
C12R1:19),  
PC (C12P21/02,C12R1:19),(C12P21/02,C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No; Location/Qualifiers  
FH key  
FT mat\_peptide 1..522  
/product='human colony-stimulating factor FT  
derivative which  
is named 'hg-CSF[ND28]';  
FT Location/Qualifiers  
FEATURES  
source  
1..525  
/organism='synthetic construct'  
/db\_xref='taxon:32630'  
BASE COUNT 94 a 184 c 148 g 99 t  
ORIGIN  
Query Match 99.7%; Score 520.4; DB 6; Length 525;  
Best Local Similarity 99.8%; Pred. No. 8.8e-87;  
Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCGCAACATATCGCGCTCGAGTCTACACAGAGCTTCTTTTAAAGAGCTTAGAGCAA 60  
DB 1 GCACCAACATATCGCGCTCGAGTCTACACAGAGCTTCTTTTAAAGAGCTTAGAGCAA 60  
QY 61 GTGAGGAAGATCCAGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGCCACCTACAAG 120  
DB 61 GTGAGGAAGATCCAGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGCCACCTACAAG 120  
QY 121 CTGTGCCACCCCGAGGAGCTGTGTCTCGGACACTCTTGGGCATCCCTGGGCTCCC 180  
DB 121 CTGTGCCACCCCGAGGAGCTGTGTCTCGGACACTCTTGGGCATCCCTGGGCTCCC 180  
QY 181 CTGAGCAGCTGCCCGAGGCGCTGCAGCTGGCAGCTGTGAGCCAACTCCATAGC 240  
DB 181 CTGAGCAGCTGCCCGAGGCGCTGCAGCTGGCAGCTGTGAGCCAACTCCATAGC 240  
QY 241 GGCCTTTTCTTACAGGGGCTCTGAGGGCTTGAAGGATCTCCCGAGTTGGGT 300  
DB 241 GGCCTTTTCTTACAGGGGCTCTGAGGGCTTGAAGGATCTCCCGAGTTGGGT 300  
QY 301 CCCACCTTGGACACTGACAGTGGAGTGGCGACTTTGCCACCACTCTGGCAGCAG 360  
DB 301 CCCACCTTGGACACTGACAGTGGAGTGGCGACTTTGCCACCACTCTGGCAGCAG 360  
QY 361 ATGGAAGAACTGGGAATGGCCCTGCAGCTGCCAGCCACCCAGGCTGCCATGCCGCTTC 420  
DB 361 ATGGAAGAACTGGGAATGGCCCTGCAGCTGCCAGCCACCCAGGCTGCCATGCCGCTTC 420  
QY 421 GCCTCTGCTTTCAGCGCGCGGAGAGGGTCTAGTCCCTCCATCTGCGAGCTTC 480  
DB 421 GCCTCTGCTTTCAGCGCGCGGAGAGGGTCTAGTCCCTCCATCTGCGAGCTTC 480  
QY 481 CTGAGGAGTCTACCGGCTTCTACGCCACTTGGCCAGGCC 522  
DB 481 CTGAGGAGTCTACCGGCTTCTACGCCACTTGGCCAGGCC 522  
RESULT 2  
LOCUS E02575 525 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA encoding human colony-stimulating factor derivative.  
ACCESSION E02575  
VERSION E02575.1 GI:2170805  
KEYWORDS JP 1990227075-A/3.  
SOURCE synthetic construct.

ORGANISM

synthetic construct  
artificial sequence.

REFERENCE

1 (bases 1 to 525)

AUTHORS

Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Ito,S.

TITLE

NEW POLYPEPTIDE

JOURNAL

Patent: JP 1990227075-A 3 10-SEP-1990;

COMMENT

KYOWA HAKKO KOGYO CO LTD

OS

Artificial gene

PN

JP 1990227075-A/3

PD

10-SEP-1990

PF

28-SEP-1989 JP 1989253097

PR

29-SEP-1988 JP 88P 245705

PI

SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI

SATO

MORIYUKI,

PI

ITOU SEIGA

PC

C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC

PC

C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,

PC

(C12P21/02,C12R1:19),(C12P21/02,C12R1:91);

CC

strandedness: Double;

CC

topology: Linear;

CC

hypothetical: No;

CC

anti-sense: No; Location/Qualifiers

FH

key

FT

mat\_peptide 1..522

FT

/product='human colony-stimulating factor FT

FT

derivative(hg-CSF[ND28N6]) which is resistant FT

FT

to proteases',

FEATURES

Location/Qualifiers

source

1..525

BASE COUNT

94 a 183 c 148 g 100 t

ORIGIN

Query Match

98.5%; Score 514; DB 6; Length 525;

Best Local Similarity

99.0%; Pred. No. 1.3e-85;

Matches 517;

Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY

1 GCGCAACATATCGCGCTCGAGTCTACACAGAGCTTCTTTTAAAGAGCTTAGAGCAA 60

DB

1 GCACCTACATATCGCAACTCGAGTCTACACAGAGCTTCTTTTAAAGAGCTTAGAGCAA 60

QY

61 GTGAGGAAGATCCAGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGCCACCTACAAG 120

DB

61 GTGAGGAAGATCCAGGCGATGCGCAGCGCTCCAGGAGAAGCTGTGCCACCTACAAG 120

QY

121 CTGTGCCACCCCGAGGAGCTGTGTCTCGGACACTCTTGGGCATCCCTGGGCTCCC 180

DB

121 CTGTGCCACCCCGAGGAGCTGTGTCTCGGACACTCTTGGGCATCCCTGGGCTCCC 180

QY

181 CTGAGCAGCTGCCCGAGGCGCTGCAGCTGGCAGCTGTGAGCCAACTCCATAGC 240

DB

181 CTGAGCAGCTGCCCGAGGCGCTGCAGCTGGCAGCTGTGAGCCAACTCCATAGC 240

QY

241 GGCCTTTTCTTACAGGGGCTCTGAGGGCTTGAAGGATCTCCCGAGTTGGGT 300

DB

241 GGCCTTTTCTTACAGGGGCTCTGAGGGCTTGAAGGATCTCCCGAGTTGGGT 300

QY

301 CCCACCTTGGACACTGACAGTGGAGTGGCGACTTTGCCACCACTCTGGCAGCAG 360

DB

301 CCCACCTTGGACACTGACAGTGGAGTGGCGACTTTGCCACCACTCTGGCAGCAG 360

QY

361 ATGGAAGAACTGGGAATGGCCCTGCAGCTGCCAGCCACCCAGGCTGCCATGCCGCTTC 420

DB

361 ATGGAAGAACTGGGAATGGCCCTGCAGCTGCCAGCCACCCAGGCTGCCATGCCGCTTC 420

QY

421 GCCTCTGCTTTCAGCGCGCGGAGAGGGTCTAGTCCCTCCATCTGCGAGCTTC 480

DB

421 GCCTCTGCTTTCAGCGCGCGGAGAGGGTCTAGTCCCTCCATCTGCGAGCTTC 480

QY

481 CTGAGGAGTCTACCGGCTTCTACGCCACTTGGCCAGGCC 522

DB

481 CTGAGGAGTCTACCGGCTTCTACGCCACTTGGCCAGGCC 522

RESULT

2

LOCUS

E02575

DEFINITION

DNA encoding human colony-stimulating factor derivative.

ACCESSION

E02575

VERSION

E02575.1 GI:2170805

KEYWORDS

JP 1990227075-A/3.

SOURCE

synthetic construct.

Db 421 GCCTCTGCTTCCAGCGCGGCGAGGGGTCTAGTTGCTCCACTGCAGAGCTTC 480

Qy 481 CTGGAGGTGCTACCGGTTCTAGCGCACCTTCCCGAGGCC 522

Db 481 CTGGAGGTGCTACCGGTTCTAGCGCACCTTCCCGAGGCC 522

RESULT 3

LOCUS E02576 525 bp DNA linear PAT 29-SEP-1997

DEFINITION DNA encoding human colony-stimulating factor derivative.

ACCESSION E02576

VERSION E02576.1 GI:2170806

KEYWORDS JP 1990227075-A/4.

SOURCE synthetic construct.

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 525)

AUTHORS Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Ito,S.

TITLE NEW POLYPEPTIDE

JOURNAL Patent: JP 1990227075-A 4 10-SEP-1990;

COMMENT KYOWA HAKKO KOGYO CO LTD

OS Artificial gene

OC Artificial sequence; Genes.

PN JP 1990227075-A/4

PD 10-SEP-1990

PF 28-SEP-1989 JP 1989253097

PR 29-SEP-1988 JP 88P 245705

PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI

SATO MORIYUKI,

PI ITOU SEIGA

PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC

C12N15/27,

PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,

PC C12P1/19)

PC (C12P21/02,C12P1/19),(C12P21/02,C12P1/91);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC Key

FT mat\_peptide 1..522

FT /product='human colony-stimulating factor FT

FT derivative

FT (hg-CSF[ND28N145]) which is resistant to FT

FT proteases;

FEATURES

source Location/Qualifiers

i..525

/organism="synthetic construct"

/db\_xref="taxon:32630"

BASE COUNT 96 a 182 c 146 g 101 t

ORIGIN

Query Match 98.2%; Score 512.4; DB 6; Length 525;

Best Local Similarity 98.9%; Pred. No. 2.6e-85;

Matches 516; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GCGCAACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTTAAAGCTTAGAGCAA 60

Db 1 GCACCAACATATCGCGCTCGAGTCTACACAGAGCTTCCTTTTAAAGCTTAGAGCAA 60

Qy 61 GTGAGGAAGATCCAGGCGATGGCGAGCGCTCCAGAGAGCTGTGTGCCACCTACAAG 120

Db 61 GTGAGGAAGATCCAGGCGATGGCGAGCGCTCCAGAGAGCTGTGTGCCACCTACAAG 120

Qy 121 CTGTGCCACCCGAGGAGCTGGTCTGCTGGGACACACTCTCTGGGCACTCCCTGGGCTCCC 180

Db 121 CTGTGCCACCCGAGGAGCTGGTCTGCTGGGACACACTCTCTGGGCACTCCCTGGGCTCCC 180

Qy 181 CTGAGAGCTGCCCGACCGAGCGCTCTGAGGTGCTGTGAGCGGTCTGTGAGCGCAACTCCATPAGC 240

Db 181 CTGAGAGCTGCCCGAGCGCGCTGCAGCTGCAGGCTGCTTTCAGCAACTCCATAGC 240

Qy 241 GGCCTTTTCTCTTACCAGGGGCTCTGCAGGCGCTTGAAGGATCTCCCCGAGTTGGGT 300

Db 241 GGCCTTTTCTCTTACCAGGGGCTCTGCAGGCGCTTGAAGGATCTCCCCGAGTTGGGT 300

Qy 301 CCCACCTTGGACACACTGCAGCTGCAGCTGCCGACTTTCACACACCATCTGCAGCAGC 360

Db 301 CCCACCTTGGACACACTGCAGCTGCAGCTGCCGACTTTCACACACCATCTGCAGCAGC 360

Qy 361 ATGAAGAAGCTGGAAATGGCCCTGCGCTGCAGCCACCCAGGTGCAATCCCGGCTTC 420

Db 361 ATGAAGAAGCTGGAAATGGCCCTGCGCTGCAGCCACCCAGGTGCAATCCCGGCTTC 420

Qy 421 GCCTCTGCTTCCAGCGCGGCGAGGAGGGTCTTAGTTCCTCCCATCTGCAGAGCTTC 480

Db 421 GCCTCTGCTTCAATCATCGGCGAGGAGGGTCTTAGTTCCTCCCATCTGCAGAGCTTC 480

Qy 481 CTGAGGTGCTGTACCGGCTTCTACGCGACACTTTCGCCAGGCC 522

Db 481 CTGAGGTGCTGTACCGGCTTCTACGCGACACTTTCGCCAGGCC 522

RESULT 4

AR024358

LOCUS AR024358 525 bp DNA linear PAT 05-DEC-1998

DEFINITION Sequence 1 from patent US 5795968.

ACCESSION AR024358

VERSION AR024358.1 GI:3977652

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 525)

AUTHORS Kuga,F., Miyaj,H., Sato,M., Okabe,M., Morimoto,M., Itoh,S.,

Yamasaki,M., Yokoo,Y., Yamaguchi,K., Yoshida,H. and Komatsu,Y.

TITLE Polypeptide derivatives of human granulocyte colony stimulating factor

JOURNAL Patent: US 5795968-A 1 18-AUG-1998;

FEATURES

source Location/Qualifiers

i..525

/organism="unknown"

BASE COUNT 85 a 192 c 151 g 97 t

ORIGIN

Query Match 93.3%; Score 486.8; DB 6; Length 525;

Best Local Similarity 98.6%; Pred. No. 1.4e-80;

Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 25 CTACACAGAGCTTCCTTTTAAAGCTTAGAGCAAGCTTAGAGAGATCCAGGGCGATGGC 84

Db 25 CTGCCCCAGAGCTTCCTGCTCAAGTGTAGAGCAAGTGAAGAGATCCAGGGCGATGGC 84

Qy 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTTACAAGCTGTGCCACCCCGAGGAGCTGGT 144

Db 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTTACAAGCTGTGCCACCCCGAGGAGCTGGT 144

Qy 145 CTGCTCGGACACTCTCTGGGGATCCCTGGGCTCCCTGAGCAGCTGCCCGAGCAGGCC 204

Db 145 CTGCTCGGACACTCTCTGGGGATCCCTGGGCTCCCTGAGCAGCTGCCCGAGCAGGCC 204

Qy 205 CTGAGCTGCGAGGCTCTGTGAGCAACTCCATAGCGGCTTTTCCTTACCAGGGGCTC 264

Db 205 CTGAGCTGCGAGGCTCTGTGAGCAACTCCATAGCGGCTTTTCCTTACCAGGGGCTC 264

Qy 265 CTGAGCGCTTGAAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324

Db 265 CTGAGCGCTTGAAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324

Qy 325 GAGCTGCGGACCTTTGCCACCCACTCTGCGACAGATGGAAGAACTGGGAATGGCCCT 384

Db 325 GAGCTGCGGACCTTTGCCACCCACTCTGCGACAGATGGAAGAACTGGGAATGGCCCT 384

QY 385 GCCTGAGCCACCCAGGAGTGCATCGCGGCTTGCCTCTGCTTTCCAGGCGCGGCA 444  
|||||  
Db 385 GCCTGAGCCACCCAGGAGTGCATCGCGGCTTGCCTCTGCTTTCCAGGCGCGGCA 444  
|||||  
QY 445 GGAGGGTCTAGTTGGCTCCATCTGCAGAGCTTCTCGAGGTGTCGACCGGTTCTA 504  
|||||  
Db 445 GGAGGGTCTAGTTGGCTCCATCTGCAGAGCTTCTCGAGGTGTCGACCGGTTCTA 504  
|||||  
QY 505 CGCCACCTTGCAGGCC 522  
|||||  
Db 505 CGCCACCTTGCAGGCC 522  
|||||  
RESULT 5  
LOCUS AR091731 525 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 1 from patent US 5994518.  
ACCESSION AR091731  
VERSION AR091731.1 GI:10018485  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 525)  
AUTHORS Kuga,T., Miyaji,H., Sato,M., Okabe,M., Morimoto,M., Itoh,S.,  
Yamasaki,M., Yokoo,Y., Yamaguchi,K., Yoshida,H. and Komatsu,Y.  
TITLE Method of producing a polypeptide having human granulocyte colony  
stimulating factor activity  
JOURNAL Patent: US 5994518-A 1 30-NOV-1999;  
FEATURES  
source Location/Qualifiers  
BASE COUNT 85 a 192 c 151 g 97 t  
ORIGIN  
  
Query Match 93.3%; Score 486.8; DB 6; Length 525;  
Best Local Similarity 98.6%; Pred. No. 1.4e-80;  
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 25 CTACCACAGAGCTCTCTTTAAAGCTTAGAGCAAGTGGAGAGATCCAGGCGGATGCC 84  
|||||  
Db 25 CTGCCACAGAGCTCTCTGCTAGTGTCTTAGAGCAAGTGGAGAGATCCAGGCGGATGCC 84  
|||||  
QY 85 GCAGCGCTCCAGGAGAGTGTGTGCACCTACAGCTGTGCCACCCCGAGGAGCTGGTG 144  
|||||  
Db 85 GCAGCGCTCCAGGAGAGTGTGTGCACCTACAGCTGTGCCACCCCGAGGAGCTGGTG 144  
|||||  
QY 145 CTGCTGGACACTCTCTGGGCAATCCCTTGGGCTTCCCTGAGCAGTGGCCCGAGCGCC 204  
|||||  
Db 145 CTGCTGGACACTCTCTGGGCAATCCCTTGGGCTTCCCTGAGCAGTGGCCCGAGCGCC 204  
|||||  
QY 205 CTCAGCTGGCAGCTCTTGGCCACCACTTCCAGTGTGGGCTTTCTCTACCGAGGGCTC 264  
|||||  
Db 205 CTCAGCTGGCAGCTCTTGGCCACCACTTCCAGTGTGGGCTTTCTCTACCGAGGGCTC 264  
|||||  
QY 265 CTCAGCGCTCGAAGAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324  
|||||  
Db 265 CTCAGCGCTCGAAGAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324  
|||||  
QY 325 GAGCTGCGGACTTTGCCACCACTTCCAGTGTGGCAGAGATGGAAGAACTGGGAATGCCCT 384  
|||||  
Db 325 GAGCTGCGGACTTTGCCACCACTTCCAGTGTGGCAGAGATGGAAGAACTGGGAATGCCCT 384  
|||||  
QY 385 GCCTGAGCCACCCAGGAGTGCATCGCGGCTTGCCTCTGCTTTCCAGGCGCGGCA 444  
|||||  
Db 385 GCCTGAGCCACCCAGGAGTGCATCGCGGCTTGCCTCTGCTTTCCAGGCGCGGCA 444  
|||||  
QY 445 GGAGGGTCTAGTTGGCTCCATCTGCAGAGCTTCTCGAGGTGTCGACCGGTTCTA 504  
|||||  
Db 445 GGAGGGTCTAGTTGGCTCCATCTGCAGAGCTTCTCGAGGTGTCGACCGGTTCTA 504  
|||||

QY 505 CGCCACCTTGCAGGCC 522  
|||||  
Db 505 CGCCACCTTGCAGGCC 522  
|||||  
RESULT 6  
LOCUS E01731 525 bp RNA linear PAT 29-SEP-1997  
DEFINITION cDNA encoding hg-CSF form human peripheral macrophage.  
ACCESSION E01731  
VERSION E01731.1 GI:2169984  
KEYWORDS JP 1988267292-A/1.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 525)  
AUTHORS Kuga,T., Komatsu,Y., Miyaji,H., Sato,M., Okabe,M., Morimoto,M.,  
Itoh,S., Yamazaki,M., Yokoo,Y. and Yamaguchi,K.  
TITLE NOVEL POLYPEPTIDE  
JOURNAL Patent: JP 1988267292-A 1 04-NOV-1988;  
COMMENT KYOWA HAKKO KOGYO CO LTD  
OS Human  
PN JP 1988267292-A/1  
PD 04-NOV-1988  
PF 23-DEC-1987 JP 1987326384  
PR 23-DEC-1986 JP 86P 306799  
PI KUGA TEISUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI, PI  
OKABE MASAMI.  
PI MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIIHARU,  
PI YAMAGUCHI KAZUO  
PC C12P21/02,C07K13/00,C12N1/20,C12N15/00//A61K37/02,A61K37/02,  
PC A61K37/02,  
PC (C12P21/02,C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC \*source: cell\_line=Peripheral macrophage;  
FH \*key Location/Qualifiers  
FH mat\_peptide 1..525  
FT /product=mature peptide of hg-CSF'.  
FEATURES  
source Location/Qualifiers  
BASE COUNT 85 a 192 c 151 g 97 t  
ORIGIN  
  
Query Match 93.3%; Score 486.8; DB 6; Length 525;  
Best Local Similarity 98.6%; Pred. No. 1.4e-80;  
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 25 CTACCACAGAGCTCTCTTTAAAGCTTAGAGCAAGTGGAGAGATCCAGGCGGATGCC 84  
|||||  
Db 25 CTGCCACAGAGCTCTCTGCTAGTGTCTTAGAGCAAGTGGAGAGATCCAGGCGGATGCC 84  
|||||  
QY 85 GCAGCGCTCCAGGAGAGTGTGTGCACCTACAGCTGTGCCACCCCGAGGAGCTGGTG 144  
|||||  
Db 85 GCAGCGCTCCAGGAGAGTGTGTGCACCTACAGCTGTGCCACCCCGAGGAGCTGGTG 144  
|||||  
QY 145 CTGCTCGGACACTCTCTGGGCAATCCCTTGGGCTTCCCTGAGCAGTGGCCCGAGCGCC 204  
|||||  
Db 145 CTGCTCGGACACTCTCTGGGCAATCCCTTGGGCTTCCCTGAGCAGTGGCCCGAGCGCC 204  
|||||  
QY 205 CTGAGCTGGCAGCTGTGTAGCCAACTCCATAGGGGCTTTCTCTACCGAGGGCTC 264  
|||||  
Db 205 CTGAGCTGGCAGCTGTGTAGCCAACTCCATAGGGGCTTTCTCTACCGAGGGCTC 264  
|||||  
QY 265 CTGAGGCGCTGGAAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324  
|||||  
Db 265 CTGAGGCGCTGGAAGGATCTCCCGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324  
|||||

```
QY 325 GACGCGCCGACTTTGGCCACCACCATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCT 384
|
|
|
Db 325 GACGCGCCGACTTTGGCCACCACCATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCT 384
|
|
|
QY 385 GCGCTGAGAGCCACCCAGGGTGCCATCGCGGCCCTTCGCTCTGCTTTCCAGCGCGGCA 444
|
|
|
Db 385 GCGCTGAGAGCCACCCAGGGTGCCATCGCGGCCCTTCGCTCTGCTTTCCAGCGCGGCA 444
|
|
|
QY 445 GGAGGGTCCCTAGTTCCTCCCATCTGCAGAGCTTCCTGGAGGTCGTACCGGCTTCTA 504
|
|
|
Db 445 GGAGGGTCCCTAGTTCCTCCCATCTGCAGAGCTTCCTGGAGGTCGTACCGGCTTCTA 504
|
|
|
QY 505 CGCCACCTTGCCCGAGCCC 522
|
|
|
Db 505 CGCCACCTTGCCCGAGCCC 522
|
|
|
RESULT 7
LOCUS E02573 525 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human colony-stimulating factor(hg-CSF).
ACCESSION E02573
VERSION E02573.1 GI:2170803
KEYWORDS JP 1990227075-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 525)
AUTHORS Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Itou,S.
TITLE NEW POLYPEPTIDE
JOURNAL Patent: JP 1990227075-A/1.
COMMENT KYOWA HAKKO KOGYO CO LTD
OS Homo sapiens (human)
PN JP 1990227075-A/1
PD 10-SEP-1990
PF 28-SEP-1989 JP 1989253097
PR 29-SEP-1988 JP 88P 245705
PI SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI
SATO MORIYUKI,
PI ITOU SEIGA
PC C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC
C12N15/27,
PC C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,
PC C12R1:19),
PC (C12P21/02,C12R1:19),(C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: cell_type=lymphoblast LUKII;
CC *source: clone=pcsf1-2;
FH Key Location/Qualifiers
FH mat_peptide 1..522
FH factor(hg-CSF);
FH Location/Qualifiers
FEATURES
source
1..525
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 85 a 192 c 151 g 97 t
ORIGIN
Query Match 93.3%; Score 486.8; DB 6; Length 525;
Best Local Similarity 98.6%; Pred. No. 1.4e-80;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 25 CTACCACAGACTTCCTTTTAAAGCTTAGACCAAGTGGAGAGATCCAGGGCGATGGC 84
|
|
|
Db 25 CTGCCCCAGAGCTTCCTGCTCAAGCTTATAGACCAAGTGGAGAGATCCAGGGCGATGGC 84
|
|
|
```

```
QY 85 GCAGCGCTCCAGGAGAAAGCTGTGTGCCACCTTACAGCTGTGCCACCCCGAGGAGCTGTG 144
|
|
|
Db 85 GCAGCGCTCCAGGAGAAAGCTGTGTGCCACCTTACAGCTGTGCCACCCCGAGGAGCTGTG 144
|
|
|
QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 204
|
|
|
Db 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 204
|
|
|
QY 205 CTGAGCTGCGAGCTCTGTGAGCCAACTCCATAGCGGCTTTTCTCTTACCAGGGGCTC 264
|
|
|
Db 205 CTGAGCTGCGAGCTCTGTGAGCCAACTCCATAGCGGCTTTTCTCTTACCAGGGGCTC 264
|
|
|
QY 265 CTGAGCGCTTGAAGGATCTCCCGAGTTGGTGTCCACCTTGGACACACTGCAGCTG 324
|
|
|
Db 265 CTGAGCGCTTGAAGGATCTCCCGAGTTGGTGTCCACCTTGGACACACTGCAGCTG 324
|
|
|
QY 325 GACGTCGCCGACTTTGGCCACCACCATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCT 384
|
|
|
Db 325 GACGTCGCCGACTTTGGCCACCACCATCTGGCAGCAGATGGAAGAACTGGGAATGGCCCT 384
|
|
|
QY 385 GCGCTGAGAGCCACCCAGGGTGCCATCGCGGCCCTTCGCTCTGCTTTCCAGCGCGGCA 444
|
|
|
Db 385 GCGCTGAGAGCCACCCAGGGTGCCATCGCGGCCCTTCGCTCTGCTTTCCAGCGCGGCA 444
|
|
|
QY 445 GGAGGGTCCCTAGTTCCTCCCATCTGCAGAGCTTCCTGGAGGTCGTACCGGCTTCTA 504
|
|
|
Db 445 GGAGGGTCCCTAGTTCCTCCCATCTGCAGAGCTTCCTGGAGGTCGTACCGGCTTCTA 504
|
|
|
QY 505 CGCCACCTTGCCCGAGCCC 522
|
|
|
Db 505 CGCCACCTTGCCCGAGCCC 522
|
|
|
RESULT 8
LOCUS E07164
DEFINITION cDNA encoding human G-CSF.
ACCESSION E07164
VERSION E07164.1 GI:2175311
KEYWORDS JP 1994092994-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 525)
AUTHORS Kuga,T., Komatsu,Y., Miyaji,H., Sato,M., Okabe,M., Morimoto,M.,
Itou,S., Yamazaki,M., Yokoo,Y. and Yamaguchi,K.
TITLE NEW POLYPEPTIDE
JOURNAL Patent: JP 1994092994-A 1 05-APR-1994;
KYOWA HAKKO KOGYO CO LTD
OS Homo sapiens (human)
PN JP 1994092994-A/1
PD 05-APR-1994
PF 23-DEC-1987 JP 1992214376
PR 23-DEC-1986 JP 86P 306799
PI KUGA NETSUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI, PI
OKABE MASAMI,
PI MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIHARU,
PI YAMAGUCHI KAZUO
PC C07K13/00,A61K37/02,C12N1/21,C12N15/27,C12P21/02,(C12N1/21, PC
C12R1:19),
PC (C12P21/02,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source
1..525
/organism="Homo sapiens"
/cell_type="macrophage"
/clone="pcsf1-2",
FT
```





```

XX OS Homo sapiens (Human)
CC PN JP 1995149798-A/1
CC PD 13-JUN-1995
CC PF 08-AUG-1994 JP 1994185787
CC PR 23-DEC-1986 JP 86P 306799
CC PI KUGA TETSUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI,
CC PI OKABE MASAMI,
CC PI MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIHARU,
CC PI YAMAGUCHI KAZUO
CC PC C07K14/535, C12N1/21, C12N15/09, C12P21/02, A61K38/00, C12N1/21,
CC PC C12R1:19,
CC PC C12P21/02, C12R1:19;
CC CC strandedness: Single;
CC CC topology: Linear;
CC FH Key Location/Qualifiers
CC FT source 1. .525
CC FT /organism="Homo sapiens"
XX Key Location/Qualifiers
FH source 1. .525
FH /db_xref="taxon:9606"
FT /organism="Homo sapiens"
XX Sequence 525 BP; 85 A; 192 C; 151 G; 97 T; 0 other;

Query Match 93.3%; Score 486.8; DB 23; Length 525;
Best Local Similarity 98.6%; Pred. No. 1.4e-80;
Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACCAGAGCTTCCTTTTAAAGCTTAGACCAAGTGGAGAGATCCAGGCGGATGCG 84
Db 25 CTGCCCCAGAGCTTCCTGCTCAAGTCTTAGACCAAGTGGAGAGATCCAGGCGGATGCG 84
QY 85 GCAGCGCTCCAGAGAGCTGTGCGCACCTACAAGCTGTGCCACCCGAGGAGCTGGTG 144
Db 85 GCAGCGCTCCAGAGAGCTGTGCGCACCTACAAGCTGTGCCACCCGAGGAGCTGGTG 144
QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTCAGCAGCTGCCCCAGCGAGCC 204
Db 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTCAGCAGCTGCCCCAGCGAGCC 204
QY 205 CTGAGCTGGCAGGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACGGGGCTC 264
Db 205 CTGAGCTGGCAGGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACGGGGCTC 264
QY 265 CTGAGGCTCCGAGAGGATCTCCCGGAGTTGGGTCCACCTTGACACACTGCAGCTG 324
Db 265 CTGAGGCTCCGAGAGGATCTCCCGGAGTTGGGTCCACCTTGACACACTGCAGCTG 324
QY 325 GAGCTCGCCGACTTTCGCCACCACTCTGGCAGCAGATGGAGAACTGGGAATGCCCT 384
Db 325 GAGCTCGCCGACTTTCGCCACCACTCTGGCAGCAGATGGAGAACTGGGAATGCCCT 384
QY 385 GCGCTCGAGCCCAACAGGGTGCCATGCGGCTTCCTGCTCTGCTTTCCAGCGCCGGCA 444
Db 385 GCGCTCGAGCCCAACAGGGTGCCATGCGGCTTCCTGCTCTGCTTTCCAGCGCCGGCA 444
QY 445 GGAGGGTCTTAGTTCCTCCATCTGCAGAGCTTCCTGGAGGTGCTGACGGCGTTCTA 504
Db 445 GGAGGGTCTTAGTTCCTCCATCTGCAGAGCTTCCTGGAGGTGCTGACGGCGTTCTA 504
QY 505 CGCCACTTTGCCAGGCC 522
Db 505 CGCCACTTTGCCAGGCC 522

RESULT 13
HUMGCSF HUMGCSF 660 bp mRNA linear PRI 27-APR-1993
LOCUS
```

## DEFINITION

Human granulocyte colony stimulating factor mRNA, complete cds.

## ACCESSION

M17706

## VERSION

M17706.1

## KEYWORDS

granulocyte colony stimulating factor.

## SOURCE

Human MIA Paca-2 cell line, cDNA to mRNA, (library of Kawasaki et al.), clone pP12.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Devlin, J.J., Devlin, P.E., Myambo, K., Lilly, M.B., Rado, T.A. and

Warren, M.K.

Expression of granulocyte colony-stimulating factor by human cell

lines

J. Leukoc. Biol. 41, 302-306 (1987)

## FEATURES

## source

Location/Qualifiers

1. .660

/organism="Homo sapiens"

/db\_xref="taxon:9606"

43. .132

/note="granulocyte colony stimulating factor signal peptide"

43. .657

/note="granulocyte colony stimulating factor precursor"

/codon\_start=1

/protein\_id="AAA35882.1"

/db\_xref="GI:183041"

/translation="MAGPQSPMKLMALQLLHSLATVQEPATPLGPASSLPQSLF

LKLEQVRKIQGDGALQKLCATYKLCHEPELVLLGSLGIPWAPLSSCPSSALQLA

GCLQSLGSLFLYQGLQALGELSPGLTDLTLDVADFPATTIQQMEELGMAPAL

QPTQGMFAFAFAQRRAGVILVASHLQSFLEVSRYLRHLAQP"

133. .654

/product="granulocyte colony-stimulating factor"

BASE COUNT 116 a 243 c 188 g 113 t

ORIGIN Unreported.

## Query Match

Best Local Similarity 93.3%; Score 486.8; DB 9; Length 660;

Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACCAGAGCTTCCTTTTAAAGCTTAGACCAAGTGGAGAGATCCAGGCGGATGCG 84

Db 157 CTGCCCCAGAGCTTCCTGCTCAAGTCTTAGACCAAGTGGAGAGATCCAGGCGGATGCG 216

QY 85 GCAGCGCTCCAGAGAGCTGTGCGCACCTACAAGCTGTGCCACCCGAGGAGCTGGTG 144

Db 217 GCAGCGCTCCAGAGAGCTGTGCGCACCTACAAGCTGTGCCACCCGAGGAGCTGGTG 276

QY 145 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTCAGCAGCTGCCCCAGCGAGCC 204

Db 277 CTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCCTCAGCAGCTGCCCCAGCGAGCC 336

QY 205 CTGAGCTGGCAGGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACGGGGCTC 264

Db 337 CTGAGCTGGCAGGCTGTGAGCCAACTCCATAGCGGCTTTTCTCTACACGGGGCTC 396

QY 265 CTGAGGCTCCGAGAGGATCTCCCGGAGTTGGGTCCACCTTGACACACTGCAGCTG 324

Db 397 CTGAGGCTCCGAGAGGATCTCCCGGAGTTGGGTCCACCTTGACACACTGCAGCTG 456

QY 325 GAGCTCGCCGACTTTCGCCACCACTCTGGCAGCAGATGGAGAACTGGGAATGCCCT 384

Db 457 GAGCTCGCCGACTTTCGCCACCACTCTGGCAGCAGATGGAGAACTGGGAATGCCCT 516

QY 385 GCGCTCGAGCCCAACAGGGTGCCATGCGGCTTCCTGCTCTGCTTTCCAGCGCCGGCA 444

Db 517 GCGCTCGAGCCCAACAGGGTGCCATGCGGCTTCCTGCTCTGCTTTCCAGCGCCGGCA 576

QY 445 GGAGGGTCTTAGTTCCTCCATCTGCAGAGCTTCCTGGAGGTGCTGACGGCGTTCTA 504

Db 577 GGAGGGTCTTAGTTCCTCCATCTGCAGAGCTTCCTGGAGGTGCTGACGGCGTTCTA 636

QY 505 CGCCACCTTGCCAGGCC 522  
 |||||  
 Db 637 CGCCACCTTGCCAGGCC 654

RESULT 14  
 HSGCSFRI 1498 bp mRNA linear PRI 21-MAR-1995  
 LOCUS Human mRNA for granulocyte colony-stimulating factor (G-CSF)  
 (PBRV-2).  
 ACCESSION X03655  
 VERSION X03655.1 GI:31693  
 KEYWORDS colony stimulating factor; signal peptide.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1498)  
 AUTHORS Nagata, S., Tsuchiya, M., Asano, S., Yamamoto, O., Hirata, Y.,  
 Kubota, N., Oheda, M., Nomura, H. and Yamazaki, T.  
 TITLE The chromosomal gene structure and two mRNAs for human granulocyte  
 colony-stimulating factor  
 JOURNAL EMBO J. 5 (3), 575-581 (1986)  
 MEDLINE 86220137  
 COMMENT Data kindly reviewed (19-JUN-1986) by S. Nagata.  
 FEATURES  
 source  
 1..1498  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 31..120  
 31..645  
 /codon\_start=1  
 /product="G-CSF protein"  
 /protein\_id="CAA27290.1"  
 /db\_xref="GI:31693"  
 /db\_xref="SWISS-PROT:P09919"  
 /translation="MAGPATQSPKMLALQLLIHWSALMTVOEATPLGPASSLPQSFL  
 LKCLEQVRKIQGDGAALQKLCATYKLCHEPELVLLGSLGIPWPLSPSCPSQALOLA  
 GCLSQLHSLFLYQGLLOALEGISPELGTLDLTLDVADFAATTIQWQEEELGMAL  
 QPTQGAMPAPAFASQRRAGGLVLAHSLQFLEVSRYRVLRLAQP"  
 121..642  
 mat\_peptide  
 /product="mature G-CSF protein (aa 1-174)"  
 1481..1486  
 misc\_feature  
 /note="polyadenylation signal"  
 1498  
 polyA\_site  
 /note="polyadenylation site"  
 281 a 488 c 402 g 327 t  
 BASE COUNT  
 ORIGIN

Query Match 93.3%; Score 486.8; DB 9; Length 1498;  
 Best Local Similarity 98.6%; Pred. No. 1.2e-80;  
 Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 25 CTACACAGAGCTCTTTTAAAGCTTAGAGCAAGTAGGAGATCCAGGCGATGGC 84  
 |||||  
 Db 145 CTGCCCCAGAGCTCTCTGCTCAAGTGTAGAGCAAGTAGGAGATCCAGGCGATGGC 204  
 |||||  
 QY 85 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAGCTGTGCCACCCGAGGAGCTGGT 144  
 |||||  
 Db 205 GCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAGCTGTGCCACCCGAGGAGCTGGT 264  
 |||||  
 QY 145 CTGCTGGACACACTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 204  
 |||||  
 Db 265 CTGCTGGACACACTCTGGGCATCCCTGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCC 324  
 |||||  
 QY 205 CTCAGCTGGCAGGCTCTTTGAGCAACTCCATAGCGGCGCTTTTCTCTACCGGGGCTC 264  
 |||||  
 Db 325 CTCAGCTGGCAGGCTCTTTGAGCAACTCCATAGCGGCGCTTTTCTCTACCGGGGCTC 384  
 |||||  
 QY 265 CTGAGCGCTTGAGGAGTCTCCCGGAGTTGGGTGCCACCTTGGACACACTGCAGCTG 324  
 |||||

Db 385 CTGACAGGCCCTGGAAGGATCTCCCGAGTTGGTCCACCTTGGACACACTGCAGCTG 444  
 |||||  
 QY 325 GACGTGCGCGACTTTTGCACACCATCTGGCAGCAGATGGAAGAACTGGGAATGCCCT 384  
 |||||  
 Db 445 GACGTGCGCGACTTTTGCACACCATCTGGCAGCAGATGGAAGAACTGGGAATGCCCT 504  
 |||||  
 QY 385 GCCCTGCAGGCCACCCAGGTTGCCATGCCGCGCTTCGCTCTTCCAGCGCGGSCA 444  
 |||||  
 Db 505 GCCCTGCAGGCCACCCAGGTTGCCATGCCGCGCTTCGCTCTTCCAGCGCGGSCA 564  
 |||||  
 QY 445 GGAGGGTCTTAGTTGCTCCATCTGCAGAGCTTCTCGAGGTGTGACCGGGTCTTA 504  
 |||||  
 Db 565 GGAGGGTCTTAGTTGCTCCATCTGCAGAGCTTCTCGAGGTGTGACCGGGTCTTA 624  
 |||||  
 QY 505 CGCCACCTTGCCAGGCC 522  
 |||||  
 Db 625 CGCCACCTTGCCAGGCC 642

RESULT 15  
 E01219  
 LOCUS cDNA encoding human G-CSF.  
 DEFINITION E01219  
 ACCESSION E01219  
 VERSION E01219.1 GI:2169478  
 KEYWORDS JP 1987129298-A/1.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1521)  
 AUTHORS Yamazaki, T., Yamamoto, O., Hirata, Y., Sekimori, Y. and Osada, J.  
 TITLE NOVEL POLYPEPTIDE  
 JOURNAL Patent: JP 1987129298-A 1 11-JUN-1987;  
 CHUGAI PHARMACEUT CO LTD  
 COMMENT  
 OS Human  
 FN JP 1987129298-A/1  
 PD 11-JUN-1987  
 PF 02-DEC-1985 JP 1985269455  
 PI YAMAZAKI TATSUMI, YAMAMOTO OSAMI, HIRATA YUICHI, PI SEKIMORI  
 YASUO,  
 FI OSADA JUICHI  
 PC C07K13/00,C07H21/04,C12N15/00,C12P21/02;  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 CC \*source: cell\_type=oral cavity tumor cells;  
 CC \*source: cell\_line=CHU-2;  
 FH Key Location/Qualifiers  
 FT CDS 31..645  
 /product="human G-CSF"  
 FT sig\_peptide 31..120  
 /note="human G-CSF prewors"  
 FT mat\_peptide 121..642  
 /note="human G-CSF"  
 FT 5'UTR 1..30  
 645..1499  
 FT 3'UTR 1500..1522  
 polyA\_site 1..1522.  
 FT mRNA Location/Qualifiers  
 1..1521  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 304 a 489 c 401 g 327 t  
 BASE COUNT  
 ORIGIN

Query Match 93.3%; Score 486.8; DB 6; Length 1521;  
 Best Local Similarity 98.6%; Pred. No. 1.2e-80;  
 Matches 491; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy 25 CTACCAGAGCTTCCTTTTAAAGCTTAGACCAAGTGAGGAAGATCCAGGGCGGATGGC 84
Db 145 CTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAAGTGAGGAAGATCCAGGGCGGATGGC 204
Qy 85 GCAGCGCTCCAGGAGAGAGCTGTGTGCCACCTACAAGCTGTGCCACCCGAGAGGAGCTGGTG 144
Db 205 GCAGGGCTCCAGGAGAGAGCTGTGTGCCACCTACAAGCTGTGCCACCCGAGGAGCTGGTG 264
Qy 145 CTGCTCGGACACTCTCTGGGATGCCCTGGGCTCCCTGTAGCAGCTGCCCCAGCCAGGCC 204
Db 265 CTGCTCGGACACTCTCTGGGATGCCCTGGGCTCCCTGTAGCAGCTGCCCCAGCCAGGCC 324
Qy 205 CTGCAGCTGGCAGGCTGCTTGAGCCAACTCCATAGGGCCCTTTTCTCTACCAAGGGGCTC 264
Db 325 CTGCAGCTGGCAGGCTGCTTGAGCCAACTCCATAGGGCCCTTTTCTCTACCAAGGGGCTC 384
Qy 265 CTGCAGGGCCCTGGAGGGATCTCCCCGAGTTGGGTCCCACTTGGACACACTGCAGCTG 324
Db 385 CTGCAGGGCCCTGGAGGGATCTCCCCGAGTTGGGTCCCACTTGGACACACTGCAGCTG 444
Qy 325 GACGTCGCCGACTTTGCCACCACTATGCGACAGATGGAAGAACTGGGAATGGCCCT 384
Db 445 GACGTCGCCGACTTTGCCACCACTATGCGACAGATGGAAGAACTGGGAATGGCCCT 504
Qy 385 GCCCTGCAGCCCAACCCAGGGTGCCATGCCGGCTTGGCCCTCTGCTTCCAGCGCCGGCA 444
Db 505 GCCCTGCAGCCCAACCCAGGGTGCCATGCCGGCTTGGCCCTCTGCTTCCAGCGCCGGCA 564
Qy 445 GGAGGGTCTCTAGTTGGCTCCCATCTGCAGAGCTTCTGGAGGTGTCGTACCGCTTCTA 504
Db 565 GGAGGGTCTCTAGTTGGCTCCCATCTGCAGAGCTTCTGGAGGTGTCGTACCGCTTCTA 624
Qy 505 CGCCACCTTGCCCGAGCCC 522
Db 625 CGCCACCTTGCCCGAGCCC 642
```

Search completed: May 8, 2002, 10:02:06  
Job time: 7729 sec